PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, 1/20 // (C12N 1:20 C12R 1:465)	A1	**(11) International Publication Number:** WO 88/ 01646 **(43) International Publication Date:** 10 March 1988 (10.03.88)

(21) International Application Number: PCT/GB87/00598

(22) International Filing Date: 25 August 1987 (25.08.87)

(31) Priority Application Number: 900,428

(32) Priority Date: 26 August 1986 (26.08.86)

(33) Priority Country: US

(60) Parent Application or Grant
(63) Related by Continuation
US 900,428 (CIP)
Filed on 26 August 1986 (26.08.86)

(71) Applicant *(for all designated States except US):* ALLELIX INC. [CA/CA]; 6850 Goreway Drive, Mississauga, Ontario L4V 1P1 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* KOZLOWSKI, Maya [CA/CA]; 3 Tullamore Road, Brampton, Ontario L6N 1J6 (CA). GLASSE-DAVIES, Roger, Wayne [GB/CA]; R.R. No. 1, Limehouse, Ontario L0P 1H0 (CA).

(74) Agent: EYLES, Christopher, Thomas; Batchellor, Kirk & Eyles, 2 Pear Tree Court, Farringdon Road, London EC1R 0DS (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: UNIVERSAL SYSTEM FOR TRANSPOSON MUTAGENESIS

(57) Abstract

Universal system for inducing genetic transposition in prokaryotic or eukaryotic cells. The system is universal in that it provides a means for inducing transposition in any organism. The invention further discloses plasmid vectors capable of mediating such genetic transposition, and novel uses for transposable elements.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

TITLE OF THE INVENTION:

UNIVERSAL SYSTEM FOR TRANSPOSON MUTAGENESIS

FIELD OF THE INVENTION

The invention pertains to a process for achieving genetic transposition in eukaryotic or prokaryotic cells. The invention also pertains to plasmids capable of mediating such transposition. The invention further pertains to novel uses of transposable elements.

BACKGROUND ART

Transposable elements are double stranded DNA molecules which possess the capacity to insert themselves into other DNA molecules. The process by which a transposable element inserts itself, termed "transposition," requires a protein known as a "transposase" (Berg, D.E. et al., Bio/Technology 1:417-435 (1983); Kleckner, N., Ann. Rev. Genet. 15:341-404 (1981)).

The transposition process results in the insertion of the transposable element into a particular site in a second DNA molecule. This insertion has four significant consequences. First, the original DNA sequence of the second (recipient) DNA molecule is disrupted. This disruption extends not only to the nucleotide sequence of the second DNA molecule, but also

to the loss of the capacity to produce the functional gene product of the previously undisrupted gene. One characteristic of transposition is that it may involve any DNA sequence of the recipient DNA molecule (i.e. transposition may be random with respect to the recipient DNA sequence being disrupted). Thus, transposition provides a powerful mechanism for mutagenizing DNA sequences.

Second, since transposition results in the incorporation of new DNA into a second DNA molecule, it provides a means of introducing heterologous DNA into a particular DNA sequence, and thus provides an alternative to cloning strategies which employ extrachromosomal plasmids.

Third, the insertion of a transposon may disrupt mRNA transcription, and hence can be used to study the transcriptional control of gene expression.

Fourth, it is possible to engineer a transposable element so that its insertion into a DNA sequence can provide one with information regarding the expression and organization of the DNA sequences which flank the site of insertion. For example, it is possible to insert a gene which encodes a non-excreted protein near to the end of a transposable element. Such a transposable element provides a probe for promoters, and secretion signal sequences (Casadaban, M., et al., Proc. Natl. Acad. Sci (USA), 76:4530-4533 (1979); Mansil, et al., Proc. Natl. Acad. Sci (USA), 82:8129-8133 (1975)). The insertion of such transposable elements places a gene whose expression can be readily monitored near to the junction between the transposable element and the disrupted DNA molecule.

Since the gene lacks a promoter region it cannot be expressed unless the insertion of the transposable element is such that it causes a promoter region, present on the disrupted molecule to become operably linked to the gene. Hence expression of the gene indicates the existence and location of a foreign promoter. The secretion of the gene product from the cell indicates the presence of a foreign secretion signal sequence at the junction site. By varying the culturing conditions it is possible to identify transposable elements which express the gene product only under certain circumstances (such as, for example, at elevated temperatures, in the presence of a particular molecule, etc.). The promoters which direct such expression are termed "conditional promoters." The identification of conditional promoters is a major present use of transposable elements. Thus, a fourth consequence of transposition is the ability to probe the genetic expression and organization of the recipient DNA molecule.

Transposable elements are diverse in both size and functional organization. Simple transposable elements, termed "insertion sequences," encode no functions unrelated to their own movement and are generally shorter than about 2 kb. Like all transposable elements, insertion sequences possess specialized termini which contain complementary sequences which are inverted repeats of one another. The presence of these inverted repeat sequences appears to be essential for transposition. (Cohen, S.N., Nature, 263:731-738 (1976)). Transposase enzymes are thought to mediate transposition by binding to DNA sequences at both ends of the transposable element.

"Transposons" are transposable elements which are larger than insertion sequences and which encode several gene products (such as proteins which confer cellular resistance to antibiotics or other selectable determinants), in addition to the transposase enzyme. Certain viruses, such as mu, lambda, or SV40 which are capable of integrating into chromosomal DNA may also be considered as extremely large (greater than 20 kb) and complex transposable elements (Cohen, S.N., Nature, 263:731-738 (1976); Cornelis, G., Bull. Inst. Pasteur, 80:3-60 (1982); Howe, M.M., Virol., 55:103-117 (1973); Nash, H.A., Ann. Rev. Genet. 15:143-167 (1981); Syvanen, M., Ann. Rev. Genet., 18:271-294 (1984); International Review of Cytology, Vol. 93, Reanney, D.C., Chambon, (Eds.), Acad. Press, Orlando, Fla. (1985)).

The bacterial transposon Tn5 has been widely studied (see, Berg, D.E., et al., supra which is herein incorporated by reference). Tn5 contains a unique central DNA sequence of approximately 2,600 bp. This central region contains genes which confer cellular resistance to streptomycin/spectinomycin, as well as an aminoglycoside phosphotransferase gene, whose product confers cellular resistance to antibiotics such as Kanamycin or Neomycin in bacteria, or G418, in eukaryotes (Berg, D.E., et al.). The presence of detectable determinants on transposons enormously facilitates their application to problems of molecular biology (Harayama, S., et al., J. Bacteriol., 153:408-415 (1983); Hui, I., et al., J. Bacteriol., 152:1022-1032 (1982)). These determinants greatly facilitate the detection of transposon containing

cells, permit the selection of populations of cells which contain transposons, and provide a means for maintaining pure cultures of transposon containing cells.

On either side of the central region is a 1,535 bp terminal region which contains the complementary inverted repeat sequences necessary for transposition. A single base pair mutation in the left-hand terminal region of Tn5 has resulted in the inactivation of the transposase gene present in the left-hand terminal region. Thus, the transposase gene present in the right-hand terminal region directs the synthesis of the transposase enzyme needed for the insertion of the transposable element.

Prokaryotic transposable elements have been identified in only a small number of bacterial strains. Although the large bulk of prokaryotic microorganisms are not known to contain transposable elements it is probable that many of them will ultimately be found to contain them. It is recognized that there is a need to study the genetic organization and expression as well as the identification of genes in many bacterial genera which do not have developed genetics. This is specially important for biotechnology and for industrially important bacteria. It is widely believed that the use of transposable elements would enormously facilitate the study of uncharacterized microorganisms. However, in many cases, it may be a difficult process to isolate and characterize transposable elements from a microorganism of interest. It would therefore be desirable to adapt well characterized transposable elements to function

in other genera of bacteria. The present inability to use transposons to study gene expression and organization has significantly encumbered efforts to understand and manipulate the genetics of many microorganisms. Among those bacteria for which no transposition system has yet been identified are the prototrophic bacteria, such as the Rhodospirillaes; the gliding bacteria, such as the Myxobacterales and the Cytophagales; the sheathed bacteria, such as Sphaerotilus; the budding and appendaged bacteria, such as Caulobacter; the Spirochetes, such as the Spirochaeteles; the spiral and curved bacteria such as the Spirillaceae; the gram-negative aerobic rods and cocci, such as the Azotobacteraceae, the Rhizobiaceae, the Methylomonadaceae, the Halobacteriaceae; the gram negative, facultatively anerobic rods, such as the Vibrionaceae, the Flavobacterium, and the Zymomonas; the gram-negative, anaerobic bacteria, such as the Bacteroidaceae; the gram-negative cocci and cocobacilli, such as the Neisseriaceae; the gram-negative anaerobic cocci, such as the Veillonellaceae; the gramnegative, chemolithotrophic bacteria, such as the Nitrobacteraceae, the Thiobacillus, the Siderocapsaceae; the methane producing bacteria, such as the Methanobacteriaceae; the gram positive cocci, such as the Micrococcaceae, the Streptococcaceae, and the Peptococcaceae; the endospore forming rods and cocci, such as the Bacillaceae; the gram positive, non-spore forming, rod shaped bacteria, such as the Lactobacillaceae; and the Actinomyces and related organisms such as the Coryneform group of bacteria, the Propionibacteriaceae, and the Actinomycetales such as the Actinomycetaceae, the

Mycobacteriaceae, the Actinoplanaceae, the Nocardiaceae and the Streptomycetaceae; the Rickettsias, such as the Rickettsiales, and the Chlamydiales; and the Mycoplasmas, such as the Mollicutes. Indeed, at present transposable elements are known only among the Enterobacteriaceae and the genus Pseudomonas.

Transposable elements have been found in eukaryotes (Reanney, D.C., et al. supra; Freeling, M., Ann. Rev. Plant Physiol., 35:277-298 (1984); Roeder, G.S., et al., In Mobile Genetic Elements, Shapiro, J.A. (Ed.), Acad. Press, N.Y., pp 299-328, (1983)).

Despite the existence of eukaryotic transposable elements, they have, thus far, not been found to be as useful in the study of eukaryotic genetics as prokaryotic transposable elements have been in the study of prokaryotic genetics. Natural eukaryotic transposable elements are significantly different from prokaryotic elements. Unlike all known prokaryotic elements, eukaryotic transposable elements are retrotransposons (Baltimore, D., Cell, 40:481-482 (1985)). The transposition of a retrotransposon requires an RNA intermediate. The formation of this intermediate and its ultimate transposition is believed to be dependent upon the secondary structure of the transposable element. Since the nature of this dependency has not been fully resolved, it is not, as yet, possible to predict the ability of an engineered eukaryotic transposable element to actually undergo transposition. Moreover, it appears that natural transposition in eukaryotes is extremely mutagenic to the transposable element, thus further restricting the ability of such elements to function as probes of gene expression (Boeke, J.D., et al., Cell, 40:491-500 (1985)).

Whereas prokaryotic transposable elements may be present in one or only a few copies per cell, eukaryotic cells may contain as many as 30-40 copies of a eukaryotic transposable element. These multiple copies of transposable elements significantly complicate genetic manipulations. Moreover, the presence of multiple copies of a transposable element may lead to recombinational events between two transposable elements resulting in the generation of additional mutations. Eukaryotic transposable elements are far less well characterized than prokaryotic transposable elements. Thus, it is at present not possible to engineer eukaryotic transposable elements to the same degree as can be accomplished with prokaryotic transposable elements. In general, eukaryotic transposable elements lack selectable determinants thus significantly complicating the selection and identification of cells that carry an engineered or recombinant transposable element. Thus, in summary, the major significant properties of prokaryotic transposable elements, which have rendered these transposable elements so useful, are largely absent from eukaryotic transposable elements.

Natural prokaryotic transposable elements are unable to undergo transposition in a eukaryotic cell. This inability results from the requirement that the transposase be present during the transposition event. If a natural prokaryotic transposable element was introduced into the nucleus of a eukaryotic cell it would not be able to express the transposase gene since the prokaryotic promoter regions of the element would not be recognized by the eukaryotic enzymes.

Even if the gene were transcribed, the RNA transcript would have to leave the nucleus in order to be translated into the transposase enzyme. Thus, even if the RNA was translated into protein, the protein would be in the cell's cytoplasm. Transposition could not occur unless the enzyme was present in the cell's nucleus where the transposable element and the recipient DNA molecule are located. Since bacterial transposases have no means of identifying and entering the cell's nucleus, they are unable to catalyze transposition in eukaroytic cells.

Jiminez, A., et al., (Nature, 287:869-871 (1980)) introduced the Tn5 transposable element into yeast in order to determine whether the element's aminoglycoside phosphotransferase gene could be expressed in eukaryotic and be used as a selectable determinant in eukaryotes. Although this work, later extended by Colbere-Garapin, F., et al., (J. Mol. Biol., 150:1-14 (1981)), and Webster, T.D., et al., (Gene, 26:243-252 (1983)), showed that the gene could be expressed in eukaryotes, it also showed that the unaltered Tn5 element could not undergo transposition in yeast. Thus, transposition in yeast and other eukaryotes could be observed only through the use of natural eukaryotic transposons and complex auxotrophic yeast strains (Scherer, S., et al., Proc. Natl. Acad. Sci. U.S.A., 76:4951-4955 (1979); Roeder, G.S., et al., Proc. Natl. Acad. Sci. U.S.A., 82:5428-5432 (1985)).

Although it has not yet been possible to select for transposition in a eukaryotic cell, techniques involving shuttle transposition have been developed which permit one to use prokaryotic transposable ele-

ments to achieve transposon-mediated mutagenesis of a eukaryotic cell. In such techniques, yeast or other eukaryotic DNA is cloned, and introduced into a bacterial strain, where a bacterial transposase catalyzes the transposition of a bacterial transposon into it. After transposition has occurred, the DNA is extracted from the bacteria and reintroduced into its original host cell, where it can then integrate by homologous recombination (Snyder, M., et al., Proc. Natl. Acad. Sci. U.S.A., 83:730-734 (1986); Seifert, H.F., et al., Proc. Natl. Acad. Sci. U.S.A., 83:735-739 (1986)). Since this technique requires homologous recombination, it is limited to those eukaryotes in which such recombination is very stingent and precise (as in yeast) since otherwise replacement of the wild type gene by the transposon-mutated gene will be a very rare event, and difficult to detect.

In conclusion, the prior art shows the desirability of using transposons to investigate, identify, modify, and control gene expression in eukaryotes. The prior art further shows the desirability of using transposons which, when integrated, permit the selection of recipient cells from the total cell population. Although natural eukaryotic transposable elements exist they are, in general, difficult to modify. Hence, the production of recombinant eukaryotic transposable elements, having the utility of prokaryotic elements, is not yet possible. By shuttling DNA between bacteria and a eukaryotic cell, it is possible to achieve transposition of eukaryotic DNA which has already been cloned. As discussed above, the applicability of this technique is quite limited.

No technique exists which is capable of directing such transposition directly in a eukaryotic cell. Such a technique would be extremely useful for the isolation and study of gene sequences.

The prior art also shows the desirability of using transposons to investigate, modify, and control gene expression in those prokaryotic microorganisms for which no known or suitable transposition system currently exists. The availability of a selectable transposition system for such microorganisms would greatly accelerate our understanding of their genetics and biochemistry. Since these prokaryotic microorganisms include the great bulk of economically important microorganisms, the ability to manipulate and control gene expression in these organisms is highly desirable.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the cloning strategy through which the plasmids of the present invention were derived.

Figure 2 shows the cloning strategy through which plasmid pMKK58 was constructed. Plasmid pMKK58 is an *E. coli* vector which contains an excisable DNA fragment containing a structural transposase gene.

Figure 3 shows the cloning strategy through which plasmid pYMK2 was constructed.

Figure 4 shows the functional map of plasmid pYMK2 which is an *E. coli*-yeast shuttle vector capable of being used to provide a functional transposase enzyme to yeast.

Figure 5 shows the restriction endonuclease cleavage map of plasmid pYMK3 which is an *E. coli*-yeast shuttle vector capable of being used to provide a

functional transposase enzyme, and a transposase element to yeast.

Figure 6 shows the results of hybridization experiments performed to demonstrate that the insertion of a eukaryotic transposable element can occur at different sites in the yeast chromosome.

Figure 7 shows a restriction endonuclease map of plasmid pYMK100, which contains a transposable element capable of controlling the expression of genes which are located near to the site of its insertion.

Figure 8 shows the use of an engineered transposable element to yield conditional mutations.

Figure 9 shows a restriction endonuclease map of plasmid SM1 which carries a transposable element and, therewithin, a Tn5 transposase gene behind a Streptomyces promoter;

Figure 10 shows a map or plasmid SM2 for Streptomyces transposition using a Mu-based transposable element;

Figure 11 shows a map of plasmid SM3, similar to pSM2 shown in Figure 10 but having the transposase gene and the transposable element physically separated by the plasmid;

Figure 12 and 13 show maps of plasmids YMK20 and YMK18, respectively, each of which is a deletion derivative of pYMK3 (Figure 3) and incapable of transcribing functional transposase;

Figure 14 is a map of plasmid YMK12 useful in eukaryotic transposition and having an ars replication region rather than a 2 μ replication region;

Figure 15 is a map of plasmid YMK30 which contains uas sequences to prevent histone binding in eukaryotes and incorporate a transposase gene behind an inducible promoter, and;

Figure 16 is a nucleotide sequence representing the uas containing segment at one terminus of the transposable element on YMK30 (Fig. 15).

SUMMARY OF THE INVENTION

A universal system is described, which is capable of mediating transposition in prokaryotic or eukaryotic cells. In order for genetic transposition to occur, a cell must be provided with a transposable element and a functional transposase gene. Through the development of a series of novel recombinant plasmids, the present invention provides a transposable element and a modified transposase gene which will permit transposition to occur in any organism.

The ability to induce genetic transposition permits the study and manipulation of prokaryotic or eukaryotic cells, and is therefore a significant advance of importance in Biology, Agriculture, and Medicine.

In detail, the present invention relates to a recombinant DNA molecule which comprises a transposase gene having a restriction endonuclease recognition site, such that the transposase gene is capable of being operably linked to a DNA sequence selected from the group consisting of: a heterologous promoter

region and/or a heterologous nuclear localization signal sequence.

The present invention additionally pertains to a DNA construct which comprises a transposase gene operably linked to a nuclear localization signal sequence or to a heterologous promoter region.

The present invention also relates to a transposable element which contains an exogenous DNA sequence.

The invention also pertains to a DNA molecule which comprises a transposable element and a DNA construct, the DNA construct comprising a transposase gene, the transposase gene being operably linked to a heterologous promoter region sequence and/or a heterologous nuclear localization signal sequence.

The invention also discloses a method for inducing genetic transposition in a prokaryotic cell which comprises:

(a) providing to the prokaryotic cell,
    (i) a transposable element, and
    (ii) a DNA construct which comprises a transposase gene operably linked to a heterologous promotor region, wherein the transposase gene expresses a transposase enzyme capable of recognizing the transposable element (i) and directing its transposition in the prokaryotic cell, and
(b) permitting the DNA construct (ii) to express the transposase gene and direct the transposition of the transposable element (i).

The invention further discloses a method for inducing genetic transposition in a eukaryotic cell which comprises:

(a) providing to the eukaryotic cell,

(i) a transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a nuclear localization signal sequence, the nuclear localization signal sequence being operably linked to a promoter region; the transposase gene being heterologous to the promoter region or the nuclear localization signal sequence, wherein the promoter region directs the synthesis of a transposase enzyme, the enzyme being linked to the amino acid sequence encoded by the nuclear localization signal sequence, and capable of entering the nucleus of the eukaryotic cell; the enzyme being capable of recognizing the transposable element (i) and directing its transposition in the eukaryotic cell, and

(b) permitting the DNA construct (ii) to express the transposase enzyme and direct the transposition of the transposable element (ii).

The invention also provides a method for controlling the expression of a target gene in a prokaryotic cell which comprises:

(a) providing to the prokaryotic cell,

(i) a transposable element, the transposable element containing an exogenous DNA sequence which comprises a conditional

and heterologous promoter region sequence capable of directing the transcription in the prokaryotic cell of an additional DNA sequence when the additional DNA sequence is linked to the transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a heterologous promoter region, wherein the transposase gene expresses a transposase enzyme capable of recognizing the transposable element (i) and directing its transposition in the prokaryotic cell,

(b) permitting the DNA construct (ii) to express the transposase gene and direct the transposition of the transposable element (i), and

(c) examining for the prokaryotic cell in which the transposition has occurred and in which the expression of the target gene is regulated by the conditional and heterologous promoter.

The invention further provides a method for controlling the expression of a target gene in a eukaryotic cell which comprises:

(a) providing to the eukaryotic cell,

(i) a transposable element, the transposable element containing an exogenous DNA sequence which comprises a conditional and heterologous promoter region sequence capable of directing the transcription in the eukaryotic cell of an additional DNA sequence when the additional DNA sequence is linked to the transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a nuclear localization signal sequence, the nuclear localization signal sequence being operably linked to a promoter region; the transposase gene being heterologous to the promoter region or to the nuclear localization signal sequence, wherein the promoter region directs the synthesis of a transposase enzyme, the enzyme being linked to the amino acid sequence encoded by the nuclear localization signal sequence, and capable of entering the nucleus of the eukaryotic cell; the enzyme being capable of recognizing the transposable element (i) and directing its transposition in the eukaryotic cell, and

(b) permitting the DNA construct (ii) to express the transposase enzyme and direct the transposition of the transposable element (i) and

(c) examining for the eukaryotic cell in which said transposition has occurred and in which the expression of said target gene is regulated by said conditional and heterologous promoter.

Additionally, the invention relates to the plasmids pMKK58, pYMK3, and pYMK100 and their functional derivatives.

Saccharomyces cerevisiae strain CMY135 (trp deletion; ura3-52) containing plasmid pYMK3 was deposited with the American Type Culture Collection (ATCC),

Rockville, MD on June 19, 1986, and designated ATCC No. 20800. E. coli. strains RR1 containing plasmids pMKK58 and pYMK100 were deposited with the American Type Culture Collection on August 1, 1986 and August 11, 1986 and given the designation ATCC No. 67173 and ATCC No. 67185 respectively.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Certain terms which are utilized in the specification and claims are defined as follows in order to provide consistent understanding thereof.

Transposition is a genetic process through which DNA sequences, termed "transposable elements," are inserted into the DNA sequence of a second DNA molecule. The process is catalyzed by an enzyme, termed a "transposase." The DNA sequence into which the transposable element is being inserted is herein termed the "recipient" DNA sequence. Importantly, transposition does not require DNA replication of the plasmid vector and thus may occur in a cell which contains a non-replicating plasmid having a transposable element and a functional transposase gene.

A "transposase" gene is one which encodes a "transposase enzyme." A transposase enzyme is capable of catalyzing the transposition of a transposable element into a recipient DNA sequence. A transposase gene is said to be capable of directing the expression or synthesis of a transposase enzyme if, upon introduction into a cell, the transposase gene provides sufficient information to permit the cell to synthesize a transposase enzyme.

In addition to the transposase gene, some transposable elements (such as, Tn3) require the expression of additional genes (such as a "Resolvase") in order to undergo transposition. Transposable elements which require such additional genes are disclosed in Mobile Genetic Elements, Shapiro, J.A., et al., eds. (Acad. Press, NY, (1983)). In employing such transposable elements in accord with the present invention, it would, of course, be necessary to additionally provide the cell with functional genes capable of expressing these enzymes.

The expression of a DNA sequence requires that the DNA sequence be "operably linked" to DNA sequences which contain transcriptional and translational regulatory information. An operable linkage is a linkage in which the regulatory DNA sequences and the DNA sequences sought to be expressed are connected in such a way as to permit gene expression. The precise nature of the regulatory regions needed for gene expression may vary from organism to organism, but shall in general include a promoter region which, in prokaryotes, contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, signal the initiation of protein synthesis. Regulatory regions in eukoytic cells will in general include a promoter region sufficient to direct the initiation of RNA synthesis.

Two DNA sequences (a promoter region sequence and a gene sequence) are said to be operably linked if the nature of the connection between the two DNA sequences does not result in either the introduction of a frame

shift mutation, or interfere with the ability of the promoter region sequence to direct the transcription of the gene sequence, or of the gene sequence to be transcribed by the promoter region sequence.

Thus, a promoter region would be operably linked to a gene if the promoter were capable of transcribing the gene. A nuclear localization signal sequence is said to be operably linked to a gene if the linkage results in the formation of a gene product which contains the amino acid sequence encoded by the nuclear localization signal sequence. A promoter region is said to be operably linked to a nuclear localization signal sequence if transcription from that promoter results in the formation of an mRNA transcript which is translatable into the amino acid sequence of the nuclear localization signal sequence.

The ability to operably link a transposase gene to a heterologous promoter region and/or nuclear localization signal sequence requires, in general, the existence of a restriction endonuclease recognition site between the transposase gene and the heterologous DNA sequence with which linkage is desired. As is well known in the art, the presence of such a site need have no effect on the transcription or translation of a DNA sequence which contains it.

Since the invention discloses the insertion of a promoter region (or nuclear localization signal sequence) into this site, any portion of the gene which originally preceded the locus of the site will be separated from the remainder of the gene by the inserted region. Thus to minimize this loss it is desirable to introduce the endonuclease recognition site

at or near the beginning of the transposase gene. The restriction site may precede the first codon of the gene but, according to the present invention, must be positioned after the end of that critical DNA sequence of the homologous promoter region which prevents the natural transposase gene from being expressed. Examples of such critical DNA sequences are the promoter sequence or the protein initiation sequence of the promoter region. Thus, for example, if a natural transposase gene could not be expressed in an organism because of the presence of a particular nucleotide or sequence of nucleotides, then it would be necessary to introduce the restriction endonuclease site after this nucleotide or sequence of nucleotides.

In general, it is not possible to know in advance the nature or location of the above described critical DNA sequences. Thus, it is preferable to introduce the restriction endonuclease site between the end of the promoter region and the beginning codons of the transposase gene, and it is most preferable to position this site between the end of the promoter region and the first codon of the transposase gene. Regardless of where this site is introduced, it is important that no DNA sequence be introduced which interfere with the transcription or translation of the transposase gene. Examples of DNA sequences which would interfere with the transcription or translation of a gene are those sequences causing frameshift or termination mutations, or those which result in the loss of the initiation codon (ATG) from the transposon gene. Thus, it is important that the introduction of the restriction site permit the transposase gene to be

operably linked to a second DNA molecule containing a heterologous promoter or nuclear localization signal sequence.

The restriction site described above may be either lost or retained after the promoter region and heterologous transposase gene have been operably linked.

The considerations inherent in the selection of the desired restriction endonuclease sequence and its location with respect to the transposase gene are well known to those of ordinary skill in the art.

The precise manner in which a transposase enzyme catalyzes transposition is not fully understood and not critical to an understanding of the present invention. Potential mechanisms for transposition are discussed in Berg, D.E., et al. supra. It is, however, widely appreciated that for transposition to occur a transposase enzyme must recognize the terminal sequences of a transposable element. A DNA sequence is said to be "recognized" by a transposase enzyme if the transposase enzyme is capable of specifically interacting with, or binding to, it.

Transposition may occur either at a specific site, a preferential site, or at a substantially random site in the recipient DNA molecule. An example of site specific transposition is the integration of the bacterial virus lambda into the E. coli chromosome (Nash, H.A., Ann. Rev. Genet., 15:143-167 (1981)). An example of transposition which results in insertion into preferential sites is the transposition of the Tn3 transposable element (Tu, C.D., et al., Cell, 19:151-160 (1980)). Examples of transposition which occur in a manner which is substantially independent

of the DNA sequence of the recipient DNA molecule (i.e. in a substantially random manner) are the transposition of the E. coli transposable element, Tn5, (Berg, D.E. et al., supra) and the integration of the E. coli virus, mu. As used herein, insertion is said to be "substantially independent" of the DNA sequence of a recipient DNA molecule if either no insertion site sequence preference is discernible, or if transposition can be observed to occur into substantially any recipient DNA sequence. As used herein, the terms "integrate," "insert," or "disrupt" are meant to be interchangable.

The term "central DNA sequence" refers to a double stranded DNA sequence which is flanked on both sides by additional DNA sequences. Those DNA sequences which flank a central DNA sequence are referred to interchangeably as either "termini sequences" or "terminal sequences." Two termini sequences of DNA are said to be "mutually complementary and inverted," if the termini sequence to the left of a central DNA sequence is capable of annealing to the termini sequence of DNA present on the right side of a central DNA sequence. An example of a DNA sequence which is mutually complementary and inverted are the DNA sequences at the termini of Tn5 (Berg, D.E. et al.).

A purely illustrative example of a transposable element is shown below:

AATGCAGGTCAGTxxxxxACTTACCTGCATT

The central DNA sequence is designated by xxxx and the mutually complementary and inverted termini sequences of DNA are represented by the flanking 13 base sequences. Although the flanking complementary and

inverted termini sequences are represented as a 13 base sequence in the above example, it is to be understood that the actual mutually complementary and inverted termini of a transposable element are considerably larger and differ from the illustrative 13 base sequences shown above. The mutually complementary and inverted termini sequences of the transposable element Tn5 are, for example, greater than 1,500 base pairs in length. The 13 base sequence shown above is provided solely to serve as an example of a mutually complementary and inverted termini sequence. As can be seen from the above example, the termini sequences need not be completely mutually complementary and inverted; a DNA sequence is considered to be mutually complementary and inverted to another DNA sequence if a substantial number of bases in the DNA sequences are mutually complementary and inverted.

According to the present invention, a transposable element may contain a "detectable marker gene" or "detectable determinant" which enables or facilitates the identification of cells which contain the transposable element. Typically, such marker genes enable the selection of cells which carry the transposable element by conferring, for example, cellular resistance to antibotics or by complementing an auxotrophic deficiency. Examples of such detectable marker genes are the aminoglyoside phosphotransferase gene of Tn5, the ura3 gene of Saccharomyces cerevisiae, or the Beta-lactamase gene of pBR322, etc.). Alternatively, a detectable marker gene may merely facilitate the screening of desired cells by expressing a product which can be readily detected. Examples of such

detectable marker genes are Beta-galactosidase, and luciferase (Engebrecht, J., et al., Cell, 32:773-781 (1983)). A transposable element, as used in the invention, may also contain additional DNA sequences within the central region of the transposable element. These additional DNA sequences, if not normally found within a transposable element, are termed "exogenous" sequences. The DNA sequence of a transposable element or a transposase gene may contain either homologous or heterologous DNA. A DNA sequence is said to be homologous to a second DNA sequence, if both sequences are normally or naturally linked to one another. A DNA sequence is said to be "heterologous" with respect to a second DNA sequence if the two DNA sequences are not normally, or naturally found to be operably linked to one another. An example of a transposable element which contains an exogenous sequence is a Tn5 transposable element whose central region carries the ura-3 gene of Saccharomyces cerevisiae. Such sequences may be either cryptic (i.e. incapable of encoding the amino acid sequence of a protein) or "expressible" (i.e. capable of encoding the amino acid sequence of a protein). Examples of a cryptic DNA sequence are: sequences which lack transcriptional or translational regulatory regions, gene fragments, and DNA sequences which contain only transcriptional and translational regulatory regions but lack transcribable DNA sequences. Examples of expressible sequences include transcribable sequences which contain transcriptional and translational regulatory regions, such as intact functional genes. It is possible to convert an expressible DNA sequence into a cryptic DNA sequence,

and vice versa, as for example by removing a promoter region from a DNA sequence which contains such a region operably linked to an intact gene, or by adding a promoter region to a DNA molecule which had previously been cryptic.

Thus, transposition requires the simultaneous presence of three elements: a transposable element, a recipient DNA molecule, and a transposase enzyme. A transposable element, as used in this invention, may be any DNA sequence which comprises a central DNA region bounded on both ends by termini regions which are recognizable by a transposase enzyme. It is possible for this central region to be of substantial size and to include several genes. It is desirable that at least one of these genes encodes a protein whose presence can be readily detected, or which confers a survival advantage to recipient cell. The central region may additionally contain a transposase gene which may either be similar to or different from the transposase gene present in the right termini region. If the product of this gene is capable of recognizing the termini DNA regions of the transposable element, then the transposable element will be capable of inserting itself into a recipient DNA molecule. If the central region does not contain a transposase gene, or if the transposase gene is not expressed or expresses a product which is incapable of recognizing the terminal DNA regions of the transposable element, then the transposable element will not undergo transposition unless a transposase enzyme capable of being expressed and functioning is additionally supplied. It is possible to direct transposition using two discrete

molecules - one containing a transposable element and the second containing a functional transposase gene.

The transposable element of the invention may be present by itself (as a discrete DNA sequence) or may be associated with cellular, viral or plasmid DNA. Typically, one associates the transposable element with cellular, viral or plasmid DNA in order to facilitate its introduction into a recipient cell. Thus, for example, one may produce a recombinant plasmid which carries the transposable element of the invention, and introduce the transposable element into a recipient cell along with the plasmid by transformation. An example of such a plasmid is pYMK3. Once introduced into a cell, the transposable element (and any associated DNA) is transported to the cell's nucleus. As used herein, the term "DNA construct" refers to a DNA sequence which has been deliberately created. A DNA construct may be present on a plasmid, virus or chromosomal DNA molecule.

A functional derivative of a plasmid is any DNA molecule which is capable of performing substantially the same functions as those which the plasmid is capable of performing.

In the present invention, a DNA sequence may be introduced into a cell by any of several means: transduction, transformation, conjugation, or microinjection, although it is most preferable to use transformation (Botstein, D., et al., The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression, Cold Spring Harbor, N.Y., 11B:607-636 (1982); Struhl, K., Nature, 305:391-397 (1983); Bollon, A.P., et al., J. Clin. Hematol. Oncol., 10:39-48

(1980); Wigler, M., et al., Proc. Natl. Acad. Sci. U.S.A., 76:1373-1376 (1979)).

One goal accomplished in the present invention is to provide a means for altering, or inactivating the expression of a target gene. A "target gene," as used herein, may refer to any recipient DNA sequence. Thus, if for example, one wished to isolate a Saccharomyces strain in which the leu-2 gene had been disrupted by transposition, then the leu-2 gene would be a target gene. Additionally, the term "target gene" may refer to any DNA sequence whose study or investigation is desired. Thus, for example, if one wished to study or investigate the leu-2 gene of Saccharomyces, then the leu-2 gene would be a target gene.

As discussed above, both the transposable element, and the recipient DNA sequence act in an essentially passive manner while undergoing transposition. The active participant in this process is the transposase enzyme (which determines, for example, the extent of insertion site specificity).

I. ENGINEERED PROKARYOTIC TRANSPOSABLE ELEMENTS

As discussed above above, a broad range of prokaryotic microorganisms cannot be studied using transposable elements because no transposition systems have yet been elucidated which are capable of functioning in these species. One aspect of the present invention is to provide means for establishing a transposition system in such prokaryotic stains.

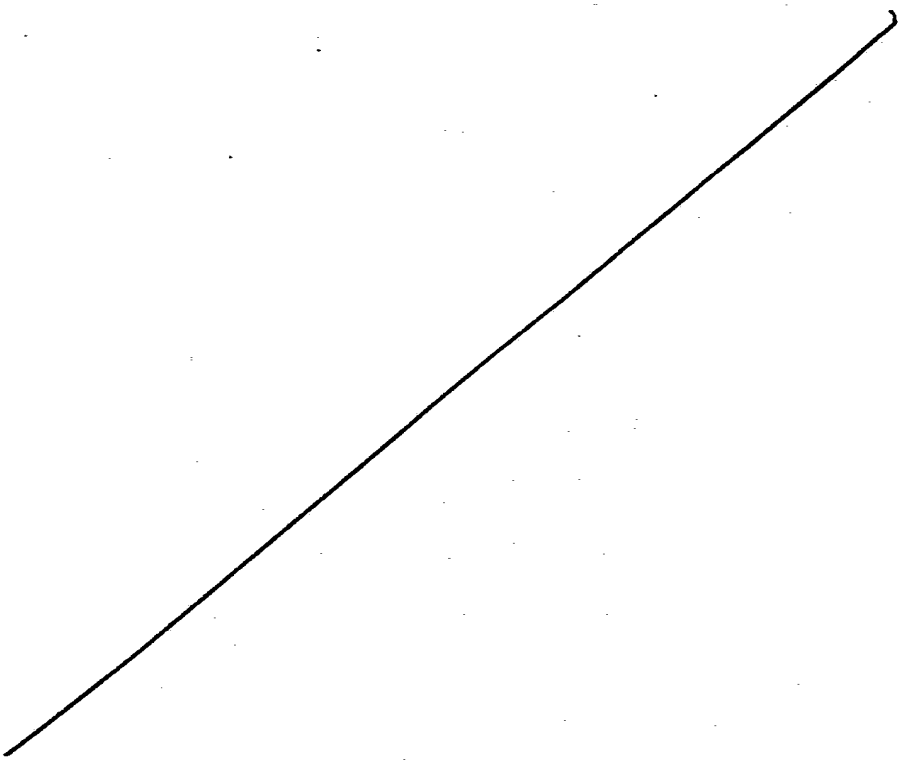
In order for a transposition system to function in a prokarytic cell, it is necessary to provide to the cell a transposable element and a functional transpos-

ase gene and, if required by the transposable element, other functional genes, such as a resolvase gene. Since the transposable element and the recipient DNA sequences are passive participants in the transposition process, a transposition system can be developed in any prokaryotic cell capable of expressing a functional transposase. The expression of a functional transposase in a prokaryotic cell, requires that a transposase gene be operably linked to a functional promoter region. The inability to observe transposition in a transformable prokaryotic cell is therefore a reflection of the inability of that cell to direct the expression of the transposase gene.

One aspect of the present invention is the discovery that this defect can be cured through the operable linkage of a promoter, endogenous to the microorganism in which transposition is sought, with a transposase gene. Thus, for example, a transposition system can be developed in any prokaryotic cell by providing to that cell the Tn5 transposable element and a Tn5 transposase which has been operably linked to an endogenous promoter of the microorganism in which transposition is sought. The present invention fulfills this need by providing a plasmid vector in which the normal Tn5 transposase promoter has been removed from the Tn5 transposase gene in such a manner as to permit any promoter region to become operably linked the Tn5 transposase gene.

Thus, one aspect of the present invention provides a system of transposition which maybe employed in any prokaryotic species. In order to establish such a transposition system, it is necessary only to operably

link a promoter from the prokaryotic species in which transposition is desired, to a functional transposase gene. When such a recombinant transposase gene is introduced into a prokaryotic cell with a transposable element, genetic transposition will occur. Thus, this embodiment of the present invention is not limited to the narrow range of prokaryotic strains in which transposition as currently been observed, but rather is a general method applicable to all prokaryotic organisms. A review of known prokaryotic promoters is provided by Hawley et al. 1983, Nucleic Acid Res, 11, 2237-2255.



II. PROKARYOTIC TRANSPOSABLE ELEMENTS CAPABLE OF UNDERGOING OR MEDIATING TRANSPOSITION IN EUKARYOTES

In order that a transposase enzyme shall be capable of identifying and entering the nucleus of a eukaryotic cell, it is necessary that the transposase gene be operably linked to a DNA sequence, termed a nuclear localization signal sequence which is itself operably linked to a functional eukaryotic promoter. A nuclear localization signal sequence is a DNA sequence which, when operably linked to a second DNA sequence and a promoter region, permits the protein specified by the DNA sequence and nuclear localization signal sequence to identify and enter the nucleus of a eukaryotic cell. Nuclear localization signal sequences are described in Kalderon, D., et al. (Cell, 39:499-509 (1984)); Silver, P.A., et al. (Proc. Natl. Acad Sci U.S.A., 81:5951-5955 (1984)); Hall, M.N., et al. (Cell, 36:1057-1065 (1984)); and Davey, J., et al. (Cell, 40:667-675 (1985)). Although any nuclear localization signal sequence may be used in this invention it is preferable to use the nuclear localiza-

tion signal sequence of the large T antigen of SV40 (Kalderon, D., et al., supra.). Any functional eukaryotic promoter, such as the Metallothionein, silk fibroin, insulin, SV_{40} Large T promoter, etc., may be employed, however it is preferable to use the ADHI promoter of Saccharomyces cerevisiae. Such linkage will result in the formation of a hybrid transposase enzyme. The nuclear localization sequence is said to be heterologous with respect to the transposase gene if the nuclear localization sequence is linked to the transposase gene by recombinant DNA techniques or is not naturally found to be linked to the transposase gene. Examples of such heterologous nuclear localization sequence transposases are: the transposase resulting from the operable linkage of an SV40 nuclear localization sequence with the Tn5 transposase gene, and a eukaryotic promoter, or the transposase resulting from the operable linkage of a transposase gene with the nuclear localization sequence of a different gene and a eukaryotic promoter.

As discussed above, in order for a transposase enzyme to be able to catalyze transposition in a eukaryotic nucleus, it must be capable of being expressed and of identifying and entering the nucleus of the eukaryotic cell. This objective is attained in the present invention through the operable (i.e., functional) linkage of a nuclear localization sequence (which is itself operably linked to a functional promoter) to a functional transposase gene. The transposase gene may be provided on a discrete DNA molecule or may, like the transposable element, be associated with cellular, viral or plasmid DNA. If the transpos-

ase gene is contained within the central region of a transposable element it will undergo transposition along with the transposable element. It is, however, possible to include a transposase gene on the same DNA molecule as that which contains the transposable element, but yet not include the transposase gene within the central DNA region of the transposable element. An example of such a DNA molecule is plasmid pYMK3 in which the transposase gene is not contained within the transposable element. In such a molecule, the transposase gene would not undergo transposition along with the transposable element.

Thus, one aspect of the present invention provides a system for establishing transposition in any eukaryotic cell. In order to establish such a transposition system it is necessary only to identify a promoter which will function in that eukaryotic cell and then to operably link such a promoter to a nuclear localization sequence which is linked to a functional transposase gene. When such a recombinant transposase gene and a transposable element are introduced into the eukaryotic cell genetic transposition will occur. Thus, the present invention is generally applicable to any eukaryote, such as the Cyanophyta, the Euglenophyta, the Chlorophyta, the Chrysophyta, the Pyrrophyta, the Phaeophyta, the Rhodophyta, the Myxomycophyta, the Eumycophyta, such as the Ascomycetes or the Basidiomycetes, the Bryophyta, the Tracheophyta, as well as the Porifera, the Mesozoa, the Coelenterata, the Ctenophora, the Platyhelminthes, the Nemertina, the Acanthocepahala, the Aschelminthes, the Entoprocta, the Ectoprocta, the Phoronida, the Brachi-

opoda, the Mollusca, the Sipunculida, the Echiurida, the Annelida, the Onychophora, the Tardigrada, the Pentastomida, the Arthropoda, the Chaetognatha, the Echinodermata, the Pogonophora, the Hemichordata, and the Chordata. Specifically included in the present invention are organisms of the phylum Chordata belonging to the classes Osteichthyes, Amphibia, Reptilia, Aves, and Mammalia.

III. USES OF THE ENGINEERED TRANSPOSABLE ELEMENT

A. Adaption of Existing Techniques

1. Engineered Transposable elements as Tools for Strain Construction

The transposable elements of the present invention may be used in the same manner as natural prokaryotic transposable elements in the formation of novel cells or microorganisms. These novel cells or microorganisms may express new desired properties or may have been altered to prevent the further expression of undesirable properties. For example, the transposable elements of the invention may be used to produce derivative strains having new combinations of genes. Alternatively, they may be used to introduce foreign or "nonnatural" DNA sequences into a cell. Thus, the transposable elements of the invention, like natural prokaryotic transposable elements, may be properly considered to be tools for strain construction.

The transposable elements of the invention may be employed as tools for strain construction in either of two possible and fundamentally different manners: by

transposition, or by genetic recombination. In the process of transposition, the transposase enzyme of the invention catalyzes the insertion of the invention's transposable element (which includes only the DNA between the termini regions of the transposable element, and no DNA beyond these regions) into the DNA sequence of a recipient DNA molecule. Concomitant with this insertion is the disruption of the gene of the recipient DNA molecule into which the insertion occurs. Thus, one result of transposition is genetic mutation.

If the insertion is substantially independent of the DNA sequence of the recipient DNA molecule, then it is possible for transposition to occur in an essentially random manner. The consequence of performing a transposition experiment on a population of cells or microorganisms, using the above-described transposase and transposable element, is the generation of a collection of derivative cells or microorganisms, each of which possesses the same transposable element, but in which the precise site of integration of the transposable element is different. Thus, by screening among these derivative cells or microorganisms, it is possible in this invention to isolate a particular cell or microorganism in which a particular gene or DNA sequence has been disrupted by the insertion of the transposable element. Since the transposable element of the invention contains a selectable marker gene, it is possible to directly select for cells which have undergone transposition.

In one embodiment of the present invention, (i.e. where the transposase gene does not undergo transposition with the transposable element) one may preclude

the possibility that multiple transposition events will occur. The transposase enzyme which is encoded by the transposase gene of the right termini region of the Tn5 is not linked to either a nuclear localization signal sequence nor a functional eukryotic promoter. Thus, this transposase gene will not be functional in a eukaryotic cell. In order for transposition to occur in such a cell, it is necessary to supply the recombinant transposase gene described above. If this recombinant transposase gene is not bounded by the terminal DNA sequences of the transposable element, then it will not undergo transposition along with the transposable element.

Since, in the above embodiment, the recombinant transposase gene does not undergo transposition with the transposable element, the transposase gene will not stably integrate or replicate in the eukroytic cell, and will be irretrievably eliminated from the cell. The resulting cell contains an integrated transposable element, but does not contain any functional transposase gene. Thus, in this embodiment a subsequent transposition event could not occur. An example of a plasmic containing a transposable element consistant with this embodiment is the plasmid pYMK3 (discussed below).

The second use of the transposable elements of the present invention, as tools of strain construction, is to significantly facilitate genetic manipulations which involve genetic recombination. In such manipulations, it is desired to either alter the expression of a particular gene, or to introduce new genetic information into a particular cell or microorganism.

Previously, such manipulations were often extremely tedious in that it was frequently difficult, or impossible, to select the desired cell or microorganism (which might be quite rare).

Transposable elements, because of their ability to confer a selectable phenotype to recipient cells, have the potential for greatly facilitating such experiments (Kleckner, N., et al., J. Mol. Biol., 116:125-159 (1977). In this use of the transposable elements of the invention, selection is initially made for insertion which occurs by transposition into a DNA sequence of a recipient DNA molecule adjacent to or within a desired target gene. The transfer of the desired target gene into a new cell or microorganism, or the mutation of the desired target gene can then be accomplished by first assaying for the presence of the transposable element (by its ability to confer a selectable phenotype), and then assaying for the expression of the desired target gene.

The transposable element of the present invention finds additional use as a portable region of DNA sequence homology. Frequently, it is desirable to combine two DNA molecules (such as, for example, two plasmids) to form a single DNA molecule. Although this goal can often be achieved through the use of restriction endonucleases, such is not always the case. For example, if the two DNA molecules being combined are uncharacterized, or contain no convenient restriction endonuclease sites, then techniques which employ restriction endonuceleases cannot be used. In such a situtation, it is possible to insert transposable elements into both DNA molecules, thereby creat-

ing regions of DNA sequence homology in both DNA molecules. Various bacterial and mammalian enzymes (such as, for example, RecA (Radding, C.M., supra.)) are capable of joining such DNA molecules through sequence specific recombination.

2. Engineered Transposable Elements as Probes of Gene Expression and Organization

Like natural prokaryotic transposable elements, the transposable elements of the present invention have significant additional uses in the fields of molecular biology and recombinant DNA technology. The insertion of the transposable elements of the present invention into a eukaryotic cell will result in a disruption of transcription. Thus by analysing either the size of RNA transcripts or the size of the peptide fragements of the disrupted gene one may determine the orientation of transcription and the location of the disrupted gene's promoter. In prokaryotic cells, the insertion of the transposable elements of the present invention result in the formation of polar mutations in the recipient DNA molecule. A polar mutation is one which results in the interruption of mRNA transcription of a DNA sequence. Thus, if a promoter is believed to direct the transcription of a DNA sequence, and the transposable element of the present invention inserts itself between the promoter site and this DNA sequence, mRNA transcription will not proceed through the transposable element. Thus, the DNA sequence which would normally have been transcribed by the promoter will not be transcribed. Hence, the

transcriptional arrangement of a gene may be investigated through the use of transposable elements. In such an investigation, one would permit transposition to occur and then assess whether a particular gene is being expressed. Failure of a particular gene to be expressed would indicate either that the transposable element had directly inserted into the target gene, or that it had inserted itself between the target gene and the normal promoter of the target gene. An example of the use of transposable elements to elucidate transcriptional patterns in bacteria is provided by Harayama, S., et al. (J. Bacteriol., 153:408-415 (1983)).

In the same manner that a natural prokayotic transposable element may be used to probe for the presence of promoters, the transposable element of the present invention may be used to search for secretion, processing, terminator, or other regulatory signal sequences.

In order to search for such signal sequences, it is preferable to employ an engineered transposable element which has been further modified for this purpose. To search for a secretory signal sequence, one would adapt the technique of Casadaban, M., et al. (Proc. Natl. Acad. Sci. U.S.A., 76:4530-4533 (1979)) and construct a transposable element which contained an exogenous gene within the left terminal repeated region. Although any gene capable of expression may be employed, it is preferable to use a gene whose expression may be easily monitored. Examples of such genes are the beta-galactosidase gene of E. coli, or the chloramphenicol acetyl transferase gene of pBR325,

etc. Regardless of which exogenous gene is employed, it is necessary to employ a genetic sequence which lacks a promoter. When such a transposable element inserts adjacent to an endogenous promoter region, the exogenous gene carried by the transposable element will be expressed, and will confer upon the recombinant cell a selectable phenotype. If the transposable element inserts next to a secretory signal sequence which is operably linked to a functioning promoter, then the resulting cell will not only synthesize the exogenous gene product, but will secrete it from the cell. If the exogenous gene product is one which can be easily monitored, its presence outside of the cell can easily be asertained.

A transposase, like many biological catalysts, is capable of catalyzing both forward and reverse reactions. Thus, a transposase may operate on a transposable element and a recipient DNA sequence to direct transposition (the forward reaction), and may operate upon an integrated transposable element to direct the excision of the transposable element from the recipient DNA molecule (the reverse reaction). The excision of a transposable element, when mediated by a transposase enzyme, is a precise event (i.e., the transposable element is excised in such a manner as to reform the exact nucleotide sequence which existed in the recipient DNA molecule prior to the insertion of the transposable element).

B. Novel Uses of the Engineered Transposable Elements

1. Using Engineered Transposable Elements to Control the Expression of an Essential Gene.

The transposable elements of the present invention, whether engineered to direct transposition in prokaryotic or eukaryotic hosts may be modified so as to enable them to identify essential genes or to regulate and control gene expression. An "essential gene" is one whose expression is necessary for cellular growth or viability. Examples of essential genes are DNA replication genes or genes which encode components of the cellular envelope.

As discussed previously, one consequence of transposition is the disruption of the DNA sequence of the recipient molecule. If a transposable element disrupts the DNA sequence of an essential gene, the essential gene will no longer be expressed and the cell will die. For this reason, it has not previously been possible to use transposable elements to identify and study essential genes. One aspect of the present invention is a transposable element which can be used to identify essential genes. To produce such a transposable element, it is desirable to modify the left-hand terminal region so as to create a functional promoter region. If the right hand terminal region is modified, the ability to express the transposase gene (which is in the right hand region) may be affected. Hence, if it is desired to modify the right hand terminal region an additional functional transposase gene may, therefore, be provided. The modification of

either termini is possible when using this invention to regulate expression in a eukaryotic cell since for such cells an additional transposase gene is generally necessary. This modification may be accomplished either by mutation, or through the cloning of a known promoter region into the transposable element. Regardless of the source or origin of the promoter region, or its position within the transposable element it is necessary according to the present invention that the promoter region direct transcription toward the terminus of the transposable element and past its end. Additionally, it is desirable that the promoter region be capable of directing conditional expression of any DNA sequence to which it becomes operably linked. Thus, for example, suitable promoter regions would be those which direct transcription only at elevated temperatures, or in the presence or absence of a biological compound (i.e. a sugar, vitamin, amino acid, etc.). Sometimes, in order to obtain such a conditional promoter region, it is necessary to clone both a promoter region and a represser or inducer gene which controls that promoter region. Examples of such promoter regions and controlling genes are the lac promoter and the lac I gene, or the pL promoter of bacteriophage lambda, and the lambda cI repressor gene. It is, however, usually possible to utilize known conditional promoter regions from the organism under study which do not require the presence of an additional repressor/inducer gene in the vector, since the host cell provides these functions. An example of such a promoter region is the gal10 promoter of Saccharomyces cerevisiae which, when used in a wild type Saccharomyces cerevisiae cell will be under the con-

trol of the endogenous products of the gal4 and gal80 genes, and require the presence of galactose in the medium for expression to occur. Figure 8 shows a diagramatical representation of this use of transposable elements.

Often, when working with poorly characterized prokaryotic or eukaryotic cells, no conditional promoter region will have been previously identified. In such a situation, as would be obvious to one of ordinary skill, the previously described procedure for isolating conditional promoters could be employed to identify such promoter regions. Once conditional promoter regions had been identified, they could be cloned and ultimately positioned into the transposable element, as described above.

When the above-described transposable elements undergo transposition in a prokaryotic or eukaryotic cell, they will, at a detectable frequency, undergo transposition next to an essential gene. If this transposition event results in the formation of an operable linkage between the essential gene and the conditional promoter region, then the essential gene will continue to be expressed despite the presence of the transposable element. Since the expression of the essential gene is dependant upon a conditional promoter, the cell will cease to grow and die when the culturing conditions are altered so as to repress or deactivate the ability of the promoter region to direct gene expression.

Thus, the presence of an essential gene, operably linked to the above described transposable element could easily be detected by allowing transposition to occur and then screening the surviving cells for those

capable of growth only under conditions which permit the expression of the conditional promoter. Because the above described transposable elements contain detectable markers, the essential gene operably linked to them may easily be isolated and subjected to further investigation. Alternatively, a transposable element containing a non-conditional promoter may be used in order to constitutively alter the expression of the adjacent gene.

2. Using Engineered Transposable Elements to Direct the Secretion of Naturally Non-Secreted Products

It is additionally possible to modify the above-described transposable elements so as to enable them to direct the secretion of normally non-secreted gene products. This can advantageously be accomplished by operably linking a secretory signal sequence to a promoter region which has been inserted into the left-hand terminal region of the transposable element. The degree and control of secretion may be accomplished by employing promoters having different levels of expression (i.e. strong or weak promoters) or by employing constitutive or conditional promoter regions.

3. Using Transposable Elements to Alter Gene Expression

The ability to operably link a conditional promoter to any gene permits one to replace or substitute a gene's normal promoter for a desired conditional promoter. Thus, if a gene is poorly expressed, or is expressed only under undesirable culturing condtions (such as subsequent to irradiation treatment, or at an

extreme temperature, etc.) it is possible to construct and isolate cells in which the desired gene is controlled by conditional promoter regions which are activated or deactivated in a more desired manner.

To summarize, the novel transposable elements of the present invention can be used to identify and study essential genes. In addition, the present invention provides a means for enabling any gene product to be expressed and excreted into the extracellular environment. Moreover, they may be used to alter the regulation and expression of any desired gene.

Having now generally described this invention, the same will be better understood by reference to certain specific examples which are included herein for purposes of illustration only and are not intended to be limiting of the invention, unless specified.

EXAMPLE 1

GENERAL PROKARYOTIC TRANSPOSITION SYSTEM

The plasmid pMKK20 was grown in E. coli strain CSH4 which contained the transposable element Tn5. Plasmid DNA was purified from this bacterial strain and used to transform RR1 cells which lacked the transposable element, and transformants which were kanamycin resistant were isolated. These transformants were found to contain plasmids which carried the Tn5 transposable element. One such plasmid, designated pMKK23 was found to contain a Tn5 element oriented so that the right hand inverted repeated termini region was adjacent to a Sal I restriction and the nucleus cleavage site. This right hand termini region was excised from the plasmid using the restriction

enzymes Sal I and Bcl I. Plasmid MP19 (P.L. Biochemicals) is a derivative of M13. Plasmid MP19 was subjected to digestion with Sal I and Bam HI restriction enzymes and the Sal I-Bcl I fragment which contained the right hand terminal region of the Tn5 was introduced into this plasmid. The Sal I-Bcl I fragment contains the Tn5 transposase gene along with its promoter region.

The MP19 derivative which contained the Tn5 transposase gene was subjected to *in vitro* mutagenesis. Techniques of *in vitro* mutagenesis involving M13 or its derivatives are disclosed by Kunkel, (*Proc. Natl. Acad. Sci. U.S.A.*, 82:488-492 (1985)) Nisbet, I.T., *et al*. (*Gene Anal. Tech.*, 2:23-29 (1985)), and Hines, J.C., *et al*., (*Gene*, 11:207-218 (1980)), which are incorporated herein by reference. In brief, the procedure entails the synthesis of a synthetic oligonucleotide having a desired and defined DNA sequence. M13, or one of its derivatives, such as MP19, is converted to its single strand form, and incubated in the presence of the synthetic oligonucleotide. Since the DNA of the oligonucleotide is controllably defined, it is possible to construct an oligonucleotide capable of pairing with a complementary DNA sequence present on the single stranded plasmid. Once base pairing has occured between the oligonucleotide and the single stranded plasmid, it is possible to extend the oligonucleotide using DNA polymerase to create a double stranded DNA molecule which may then be sealed by DNA ligase. When this double stranded DNA molecule is introduced into a bacterial cell semi-conservative DNA replication will result in the production of progeny molecules which now contain the DNA sequence of the

oligonucleotide fragment (Messing, J., et al., Nucl. Acid Res., 9:309 (1981). Thus, if one desired to introduce a point mutation, or an exogenous DNA sequence into a specific site on a plasmid one would design an oligonucleotide fragment which contained the mutation or exogenous DNA sequence and then pursue the above described procedure. In order to introduce this mutation or exogenous DNA sequence into a particular region of a plasmid, it is necessary to surround the mutation or the exogenous DNA sequence with flanking DNA sequences which are complementary to the DNA sequence of the region whose mutagenisis is desired.

An oligonucleotide of 28 units having the sequence:

5'-CAGAAGTTATCAGGGTCGACAACGTTAC-3'

was synthesized, and allowed to aneal to the single stranded MP19 plasmid which contained the Tn5 transposase gene. This oligonucleotide was used as a primer for DNA polymerase to form a double stranded DNA plasmid which was made covalently closed through the action of DNA ligase. This plasmid was transformed into E. coli strain JM103 (Messing, J., et al., supra) and a plasmid which had incorporated the oligonucleotide fragment described above was identified and isolated. This oligonucleotide contains a Sal I restriction endonuclease cleavage site which is bracketed by DNA sequences which separate the promoter region of the Tn5 transposase gene from the functional transposase gene itself.

The above procedure was successful in engineering a plasmid, designated MP19-Transposase 57 construct which contains a Sal I restriction endonuclease cleav-

age site separating the Tn5 transposase gene from the promoter of that gene (i.e. the promoter regions and the transposase gene are operably linked). The presence of this restriction endonuclease cleavage site does not affect the ability of the normal Tn5 promoter to direct the transcription of the Tn5 transposase gene. Thus, plasmid MP19-Transposase 57 construct is capable of providing a functional transposase to E coli and related genera. The cloning strategy which led to the formation of MP19-Transposase 57 construct is shown in Figure 1.

Plasmid MP19-Transposase 57 construct was incubated in the presence of the restriction endonucleases SalI and EcoRl and a 1.5 kb restriction fragment which contained the Tn5 transposase gene was isolated. This 1.5 kb fragment lacked the normal transposase promoter region. Plasmid pBR322 was incubated in the presence of the SalI and EcoRl endonucleases and then incubated with DNA lisase and the 1.5 kb transposase gene fragment. Through this procedure a plasmid, designated pMKK58 was constructed. This cloning strategy is shown in Figure 2.

Plasmid pMKK58 contains the Tn5 transposase gene. This gene is not operably linked to any promoter but is preceded by a Sal I restriction endonuclease site into which a promoter region or a nuclear localization signal sequence could be inserted to form an operable linkage with the transposase gene.

Importantly, the presence of the Sal I restriction endonculease cleavage site of plasmid pMKK58 permits one to operably link the Tn5 transposase to any promoter region from any prokaryote.

EXAMPLE 2

ESTABLISHING A SYSTEM FOR TRANSPOSITION IN A PROKARYOTIC MICROORGANISM

In order to establish a system for transposition in any prokaryotic microorganism, DNA from the microorganism in which transposition is desired is extracted and purified by means well known in the art. The purified DNA is then subjected to random restriction endonuclease digestion. Although it is preferable to use the restriction endonuclease Sal I to accomplish this digestion, any restriction endonuclease may be employed. Plasmid pMKK58 is incubated with Sal I restriction endonuclease under conditions sufficient to cleave the plasmid molecule at the Sal I recognition site present between the Tn5 transposase gene and its promoter. The random digestion fragments (some of which will contain functional promoter fragments) obtained by restriction endonuclease cleavage of the DNA of the prokaryotic microorganism is then incubated in the presence of the Sal I-digested plasmid pMKK58 under conditions sufficient to permit the incorporation of one or more of the random fragments to be incorporated into the Sal I site preceding the transposase gene. As is well known by one of ordinary skill in the art, if a restriction endonucelase other then Sal I or an isoschizomeric endonuclease is employed, it will be necessary to adjust the termini of the plasmid and random fragments so as to make them complementary and ammenable to DNA ligation. The procedures need for the performing these manipulations are set forth in Maniatis, T., et al., (Molecular Cloning a Labortary Manual, Cold Spring Harbor, NY (1982)) which is herein incorporated by reference.

The result of the above manipulation is the generation of a pool of recombinant plasmids some of which will contain a promoter region from the prokaryotic microorganism in which transposition is desired, operably linked to the Tn5 transposase gene.

The above described plasmid pool is them introduced into the prokaryotic microorganism (in which transposition is desired) along with any transposable element whose terminal DNA regions are recognized by the Tn5 transposase enzyme. Although any of the previously described methods for introducing DNA into a prokaryotic microorganism maybe employed, it is in general preferable to employ transformation. Transposition will occur at a detectable frequency in those prokaryotic microorganism which have received both a plasmid containing a transposable element and a plasmid in which the Tn5 transposase gene has been operably linked to a cellular promoter region.

EXAMPLE 3

ESTABLISHMENT OF A TRANSPOSITION SYSTEM IN A PROKARYOTIC MICROORGANISM USING PREVIOUSLY CLONED GENES

DNA from any prokaryotic microorganism may be isolated and purified by means well known in the art. Such DNA may then be subjected to restriction endonuclease cleavage and religated into a plasmid, such as, for example, pBR322, or alternatively any plasmid capable of replication in the microorganism in which transposition is desired.

Plasmids obtained by the above described procedure are screened to identify individual plasmids which contain DNA inserts which give rise to the production

of protein molecules. Such plasmids in general shall contain a cellular promoter which is operably linked to a prokaryotic gene. By techniques well known in the art, one can establish the location and direction of the transcription which proceeds from the cloned promoter region and construct plasmid vectors in which this cloned prokaryotic promoter region has been isolated from the DNA sequences with which it was originally associated.

By the above described method, it is possible to identify promoter regions from any prokaryotic microorganism. Such promoter regions may be inserted, by means well known in the art, into the Sal I restriction endonuclease recognition site of plasmid, pMKK58 and the recombinant plasmid in which the prokaryotic promoter is operably linked to the Tn5 transposase gene may then be isolated and purified.

In one embodiment, a culture of a microorganism in which transposition is desired is provided with two plasmids, one of which contains a transposable element and the second of which contains the above described plasmid having a prokaryotic promoter operably linked to the Tn5 transposase gene. If one first selects for cells which have lost the transposable element containing plasmid (by, for example, using a plasmid which cannot replicate, or by treating the transformed cells with a plasmid curing agent such as acridine orange) and then screens or selects for the detectable marker of the transposable element one would isolate prokaryotic microorganisms in which transposition had occurred. Alternatively, the invention could be performed with a single plasmid having both a trans-

posable element and the above-described transposase gene.

As would be obvious to one of ordinary skill in the art, the preceding examples could be performed using any transposable element and are not restricted to the use of Tn5. For example, the transposase genes of other transposable elements could be isolated and operably linked to a heterologous promoter region. When such a construct, and a transposable element (recognized by the transposase) are introduced into a cell capable of recognizing the heterologous promoter region, genetic transposition will result. As indicated previously, some transposable elements require the expression of additional genes in order to undergo transposition (i.e. Tn3 and Tn10 require a "resolvase"). Hence before such transposable elements will undergo transposition it will be necessary to clone these additional genes such that they are operably linked to a promoter capable of being recognized by the cell in which transposition is sought. When using a transposase gene other than that from Tn5, the restriction endonuclease cleavage site which is introduced between the endogoneous transposase promoter and the transposase gene is preferably one which is known not to cleave elsewhere in the transposase gene. Such a restriction endonuclease can easily be identified by incubating a plasmid which contains the transposase promoter-transposase gene with various endonucleases and identifing a restriction endonuclease which fails to cleave the plasmid into a linear form. The recognition site of such an enzyme then be synthesized and used as the oligonucleotide fragment as explained in Example 1.

Example 3

Transposon Mutagensis of Streptomyces

An ability to transform organisms of the genus Streptomyces by transposition would be of particular value, given the industrial importance of this genus as sources of antibiotic, anti-viral and anti-cancer compounds. For reasons given previously, transposons are particularly powerful tools for genetically analyzing such organisms as Streptomyces although, to date, no mechanism for attaining transposition has been reported for this genus.

It is generally accepted in the art, for example, that such native transposable elements as Tn5, Tn916, Tn3 and Tn10 are unable to transpose in Streptomyces. To illustrate the ability of transposable elements designed in accordance with the present invention to overcome the drawbacks associated with obtaining transposition events in Streptomyces, three transposable elements of varying design were constructed and incorporated on a Streptomyces plasmid vector. Each of the elements was constructed using the principle on which aspects of the present invention are based i.e. by replacing the natural transposase gene promoter with a promoter which, in the present example, functions in Streptomyces.

Each of the three constructed elements described below was incorporated on a cloning vector known to function in Streptomyces. This vector pMT660 (see Birch and Cullum, 1985, Microbiol., 131, 1299-1303) exhibits temperature sensitive replication, becoming unstable at 42°C. By raising experimental temperatures therefore the plasmid will be lost from the culture, facilitating detection of a transposition event. Other distinguishing features of pMT660 include resistance to thiostrepton (Ts^r) which is therefore useful as a selectable marker, and the availability of various unique restriction sites, including ClaI and KpnI.

The three engineered transposable elements components of which were assembled in pUC18 and then incorporated into pMT660 as a pUC18 fragment, are described below and illustrated in accompanying drawings as noted.

1. STN1

The transposable element STN1 is based on well known, broad host range transposon Tn5 and retains many of its valuable features including the transposase structural gene, transposase binding sites i.e. termini of Tn5, and a region which encodes kanamycin resistance (Km^r). To adapt the transposon Tn5 for use in Streptomyces however the natural promoter of the Tn5 transposase gene was replaced with a Streptomyces-recognized promoter i.e. the Pc promoter from

plasmid IJ101 (see Deng et al., Gene 43, 295-300) which promoter is known to function in both E. coli and Streptomyces.

In addition to replacing the promoter for the transposase gene, the natural promoter of the Km^r gene of Tn5 was also replaced by the Pc promoter. In general, it should be recognized that when an exogenous gene is contained on or incorporated into a transposable element, it is preferable and may be necessary to replace the promoter of that gene as well with a promoter better suited to the intended host to facilitate expression of the exogenous DNA.

As a further modification of Tn5 to produce STN1, the termini of Tn5 were truncated to provide for the minimum transposase binding sequence at the termini of the element as described, for example, by Kwoh et al., 1981, Gene 13, 37-46.

Thus STN1 fragments were individually synthesized, assembled in pUC18 and transferred to pMT660, providing pSM1 as shown in Figure 9. As a transposable element, therefore, STN1 provides truncated Tn5 termini between which is a coding region for Tn5 transposase under Pc promoter influence and, in tandem a coding region for Km^r also under the influence of its own Pc promoter.

2. STN2

The transposable element is based on the phage Mu (see Bukham, Ann. Rev. Genet. 1976, 10, 389-412), known to be a large transposable element. Phage Mu is, like Tn5, bound by sites to which transposase binds, which sites may be truncated without losing their affinity for transposase binding. The Mu transposase enzymes MuA and MuB are encoded in a region between the Mu termini and are under the influence of the natural promoter for MuA,B transposases expression. Also within the central region of phage Mu are regions encoding numerous phage proteins.

As a transposable element, phage Mu has the advantages that it has been shown to transpose in vitro indicating that host factors are not required; also, its transposase enzymes MuA,B act in trans i.e. are functional at a site remote from its site of synthesis, suggesting that it may be better able to function in a heterologous environment, such as in Streptomyces which is not a natural Mu host.

To create STN2, plasmid IF110 (see _____) which contains a Km^r gene between truncated Mu ends i.e. a "mini-Mu" segment was cut to remove the mini-Mu segment. Thereafter the Km^r fragment was removed and replaced with the Pc promoter/Km^r coding region constructed in STN1 as

described above, thereby incorporating between the Mu ends a Km^r coding region under the influence of a Streptomyces-functional promoter. Also incorporated within the Mu ends was a segment coding for MuA,B transposase under the influence of the tac promoter (see Amann et al., 1983, Gene, 25, 167-178). Again, each segment was assembled in pUC18 and then ligated into the ClaI site of pMT660, thereby creating pSM2 shown in Figure 10.

3. STN3

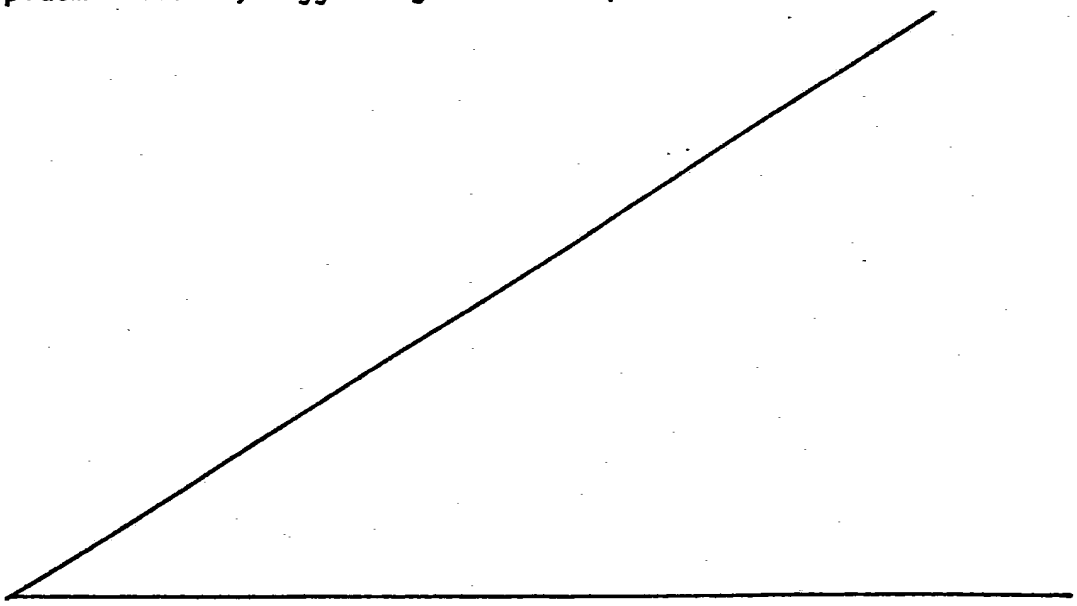
The transposable element STN3 is similar to STN2 in that both are based on the phage Mu. Element STN3 is by design however, actually two separate segments one which codes for Mu transposase and another, separated on the pMT660 vector, which codes for Km^r mini Mu. In the transposase coding segment, the MuA and B transposase coding region is placed behind the promoter of the thiostrepton resistance gene of pMT660. The mini-Mu/Km^r segment of pSM2 was incorporated on pMT660 by substituting it for the Ts^r gene of pMT660.

Thus pSM3 separates the transposase enzyme producing segment and the transposable element segment. This separation distinguishes pSM3 from pSM2 in which the two segments are combined on the transposable element STN2. Plasmid pSM3,

which incorporates these components separately is shown in Figure 11.

Transposition of Streptomyces by STN1

Streptomyces strain TK24 was transformed with pSM1 which contains transposable element STN1. Transformants were selected for growth in the presence of kanamycin and then screened for thiostrepton resistance at 30°C. The Km^r, Ts^r colonies were transferred to kanamycin containing growth medium and incubated at 42°C, a temperature at which the plasmid pMT660 is unable to replicate. Colonies were recovered which exhibited the Km^r, Ts^s phenotype, indicating that (1) the plasmid was unstable at 42°C and the Ts^r phenotype was coordinately lost and that (2) the Km^r phenotype was retained despite the absence of functional plasmid vector, suggesting that transposition had occurred.



Example 4

DEMONSTRATION OF TRANSPOSITION IN EUKARYOTES

The 1.5 kb Eco RI-Sal I fragment of plasmid PMKK58 which contains the Tn5 transposase gene was excised from plasmid PMKK58 and purified. A synthetic nuclear localization signal sequence derived from the large T antigen of SV40 (described previously) was synthesized. This oligonucleotide had the following sequence:

AATT ATG CCA AAG AAG AAG AGA AAG GTC GAAGATCCA TCGAC
(PRO LYS LYS LYS ARG LYS VAL)

The above oligonucleotide has three important properties. Firstly, the left most end provides a recognition site for the Eco RI* restriction endonuclease, whereas the right hand termini contains a recognition site for the Sal I restriction endonuclease (Maniatis, T., et al., supra). Secondly, the sequence contains an ATG codon which would be recognized in a eukaryotic cell as a translation initiation site. Lastly, the oligonucleotide contains a 21 base pair long sequence immediately adjacent to the ATG codon which duplicates the nuclear localization signal sequence of the T-antigen. The Sal I end of the synthetic oligonucleotide was ligated to the Sal I end of the 1.5 kb fragment described above thereby generating a single fragment having 2 Eco RI ends. This fragment which contains the synthetic nuclear localization signal and the Tn5 transposase gene was then introduced into the Eco RI site of plasmid pYCDE2. Plasmid pYCDE2 contains the origin of replication of the yeast 2 micron plasmid (Hollenberg,

C.P., Curr. Topics, Microbiol. Immunol., 96:119-144 (1982)) as well as the origin of replication of plasmid pBR322 (Bolivar, F., et al., Gene, 2:95-113 (1977)). Plasmid pYCDE2 also contains the ADHI promoter region of Saccharomyces cerevisiae. The Eco RI site described above is immediately adjacent to the ADHI promoter region.

The plasmid derived from this work, was designated pYMK2 and contained the ADHI promoter operably linked to a nuclear localization signal sequence which was to the Tn5 transposase gene. Thus, plasmid pYMK2 was capable of expressing a transposase eukaryotic cell which is capable of reentering the cellular nuclease. Plasmid pYMK2 is therefore a general eukaryotic vector capable of being used to provide a functional transposase enzyme to the nucleus of any eukaryotic cell which is capable of transcribing a gene from the least ADHI promoter. The cloning strategy through which plasmid pYMK2 was constructed is shown in Figure 3. The functional map of plasmid pYMK2 is shown in Figure 4. In addition, plasmid pYCDE2 contains the Saccharomyces cerevisiae trp1 gene (Department of Genetics, University of Washington, Washington Research Foundation).

A Tn5 element was modified so as to additionally contain the ura3 gene of Saccharomyces cerevisiae. In order to accomplish this, a plasmid containing the Tn5 element was purified and subjected to digestion with the endonuclease BamHI. This endonuclease recognizes and cleaves a site present in the central region of the Tn5 transposable element. A 2.2 kb BamHI-BgI II fragment which contained the ura3 gene of Saccharomyces cerevisiae (Rose, et al., Gene, 29:113-124 (1984) was ligated into the BamHI site and a plasmid

containing a Tn5 derivative which contained the ura3 gene was isolated.

Plasmid pYMK2 was transformed into an E. coli strain which contained the above described modified Tn5 transposable element. After overnight growth, the transformants were pooled and plasmid pYMK2 DNA was extracted from them and purified. This preparation of plasmid pYMK2 was transformed into an E. coli strain which did not contain Tn5. Among the resulting transformants were those which exhibited resistance to kanamycin. The plasmid from these transformants were isolated and shown to contain the Tn5 transposable element. These plasmids were screened to identify plasmids in which Tn5 had inserted into a region other than either the 2 micron origin, trp-1 gene, the beta-lactamase gene, the pBR322 origin of replication, or the recombinant transposase gene. Plasmids fulfilling these requirements are therefore suitable for providing both a transposable element and a functional transposase gene to a eukroyotic cell. When such a plasmid is transformed into a eukroyotic the recombinant transposase gene is expressed and catalyzes the transposition of the plasmid-borne Tn5 transposable element into a recipient DNA molecule present in the nuclease of the eukroytic cell. This procedure resulted in the isolation of pYMK3 which fulfilled all of the above requirements. The restriction endonuclease map of pYMK3 is shown in Figure 5.

Example 5

TRANSPOSITION OF TRANSPOSABLE ELEMENTS WHICH CONTAIN EXOGENOUS GENES IN EUKARYOTIC CELLS

Plasmid pYMK3 may be envisioned as being composed of two connected regions. The first, or "plasmid region" contains the *trp-1*$^+$, and the beta-lactamase genes as well as the both yeast and bacterial origins of replication. In addition this region also contains the hybrid transposase gene with its eukaryotic promoter. The second, or "transposable element region" contains the yeast *ura-3*$^+$ gene and the antibiotic resistance determinant gene. The entire transposable element region is bracketed by terminal DNA regions which contain the mutually complementary and inverted DNA sequences of Tn5.

Plasmid pYMK3 was introduced into *Saccharomyces cerevisiae* spheroplasts of strain CMY135 according to the method of Hinnen, A., *et al*. (Proc. Natl. Acad. Sci. U.S.A., *75*:1929-1933 (1978)). CMY135 yeast cells are deficient in their capacity to synthesize tryptophan (*trp* deletion) and uracil (*ura3-52*). The deficiencies caused by these genetic lesions may be complemented by the *trp* and *ura* genes carried by plasmid pYMK3. It is not however, possible to repair these lesions through homologous recombination with the plasmid-borne sequences.

The mixture of transformed cells and pYMK3 plasmid was plated on culture medium lacking uracil in order to select for the growth of yeast colonies from cells which had received the plasmid pYMK3. Cells were grown on *C Medium. *C Medium contains (per 500 ml): 3.35g yeast nutrient broth, 4.35g K_2HPO_4, 2.9g

succinic acid, 20 ml of 50% glucose, 16.6 ml of 60% glycerol, and 25 ml *C concentrate. *C concentrate contains (per 300 ml): 120 mg each of adenine, arginine, histidine, isoleucine, leucine, methionine, tryptophan (unless deleted), tyrosine and uracil (unless deleted); 180 mg of lysine; 300 mg of phenylalanine, 600 mg of threonine and 900 mg of valine. After 48 hours of incubation at 37°C, 2000 colonies/plate had formed, indicating a transformation frequency of 200 transformants/ug of DNA. These ura$^+$ yeast colonies were found to belong to one of two classes: ura$^+$ trp$^-$ (1%) or ura$^+$, trp$^+$ (99%). Those ura$^+$ colonies which were also observed to be proficient in the biosynthesis of tryptophan (trp$^+$) were found to contain the autonomously replicating plasmid pYMK3. Thus, in these cells the ura-3 and trp1 lesions were found to have been complemented by the ura3$^+$ and trp1 genes of plasmid pYMK3.

The detection of ura$^+$, trp$^-$ colonies indicated that either genetic transposition or an anomalous rearrangement or recombinantional event had occurred in approximately 1% of the ura$^+$ colonies. The fact that these colonies no longer suffered from the effects of the ura3-52 mutation (i.e., the colonies were ura$^+$), but still retained their auxotrophy for tryptophan (i.e., the cells were trp$^-$) indicated that in these cells the DNA sequences of the ura3 and trp1 genes (which had both been present on plasmid pYMK3) had become physically separated. Since genetic recombination between the ura3-52 mutation and the trp deletion of the yeast strain and the ura3 and trp1 genes of plasmid pYMK3 was not capable of restoring ura$^+$ or

trp+ activies the isolation of these colonies suggested that a process other than either plasmid replication or generalized recombination had occurred. The effort to purify plasmid DNA from ura+ trp- isolates and rescue it in E. coli failed. This result suggested that pYMK3 had undergone a rearrangement resulting in a loss of the pBR322 portion. Two experiments were then performed which confirmed that pBR322 DNA was no longer present in yeast cells. 1) The ura^+ trp^- isolates were not able to produce a zone of clearance on a starch-iodine medium, which is normally observed in the AP^r, p-lactamase producing cells (Chevaltier, et al. FEBS letters, 108:179-180 (1979)). 2) No homology was found when total DNA purified from the ura+ trp- cells was probed with radioactively labeled pBR322.

Separation of ura and trp markers when introduced to yeast in standard cloning vectors has never been observed. In order to substantiate that the ura^+ trp^- isolates arose as a result of an active transposase gene present on pYMK3, deletion derivatives of pYMK3 were constructed. In one case, the ADH1 promoter and about 600 bases of the 5' end of the transposase gene were removed by cleaving pYMK3 with SalI and XhoI enzymes and then religating the vector, thereby providing pYMK20 shown in Figure 12. The parallel, another deletion derivative was created by restricting pYMK3 with Hind III to remove about 400 bases from the 3' end of the transposase gene, and the plasmid religated to provide pYMK18 shown in Figure 13.

When yeast cells were separately transformed with pYMK3, pYMK18 and pYMK20, ura$^+$ trp$^-$ colonies were observed only for those cells transformed with pYMK3 but not those transformed by the deletion derivatives pYMK18 or pYMK20 which are incapable of providing a functional mRNA transcript of the transposase gene. The inability of the deletion derivatives to code for functional transposase mRNA was demonstrated by Northern analysis. Whereas RNA extracted from the host yeast CMY 135 contained a component RNA which hybridized to a transposase DNA probe when the host was transformed by plasmids pYMK2, pYMK3 and pYCDE2, no hybridizable RNA component was detected in the hosts transformed with the deletion derivatives pYMK18 or pYMK20.

Example 6

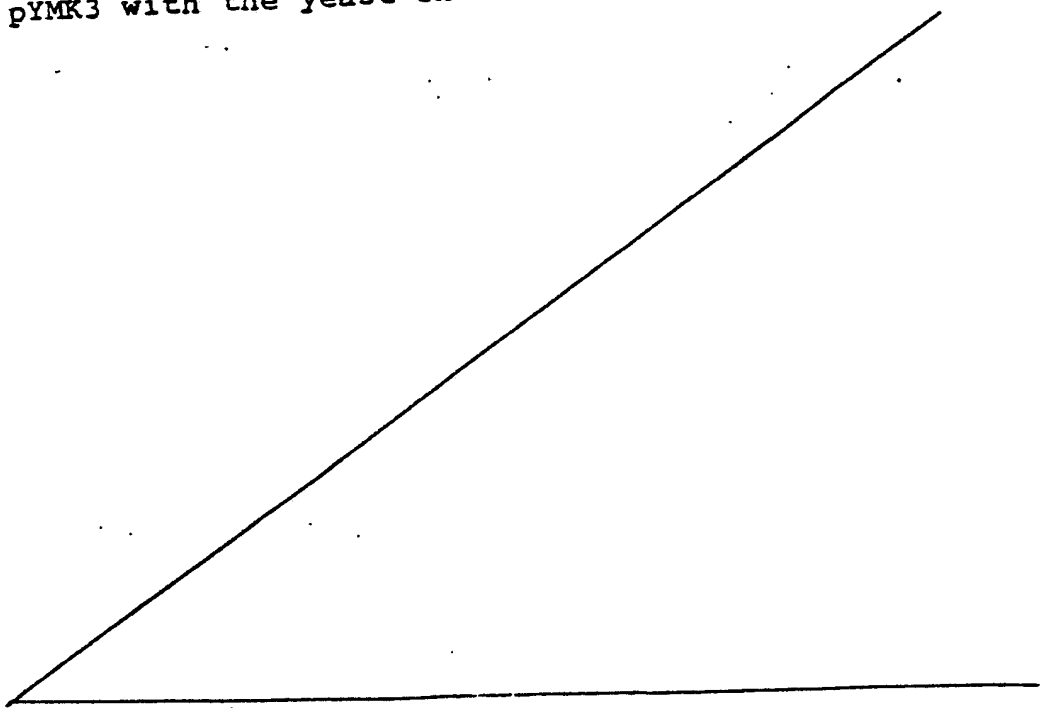
DEMONSTRATION OF THE INSERTION OF AN ENGINEERED TRANSPOSABLE ELEMENT INTO THE CHROMOSOME OF SACCHAROMYCES CEREVISIAE

In order to isolate the yeast chromosomal DNA which was linked to the inserted transposable elements of the ura$^+$ trp$^-$ yeast cells of Example 4, the total DNA of these cells was partially digested with the EcoRI restriction endonuclease. The DNA fragment obtained from this treatment were separated according to size on a 10-40% glucose gradient. Fragments of similar size were pooled, the DNA precipitated with ethanol, dried *in vaccuo* and resuspended in 100-200 ul of TE buffer (10mM Tris, 1mM EDTA pH8.0). The DNA fragments from each pool were ligated into the EcoRI site of the plasmid pBR329 (Covarrubias, *et al*. (*Gene*, 17:79-89 (1982)). *E. coli* strain RRI were transformed with the ligation mixtures. Transformants were selected on agar containing kanamycin and screened for

resistance to ampicillin (Ap^R), tetracycline (Tc^R) and kanamycin (Km^R). This procedure resulted in the identification of a group of plasmids in which the cloned yeast chromosomal DNA contained a kanamycin resistance determinant. Since the kanamycin resistance determinant was originally associated with a transposable element, the plasmids were further analyzed to determine whether they contained the ura3 gene of the transposable element. Thus, E. coli strain FB1009, which is ura^-, was transformed with these colonies to see whether the ura^- mutation could be complemented. All Km^R isolates were found to be able to complement the ura^- mutation in this strain.

The cloned yeast DNA was then analyzed with restriction endonucleases. The patterns generated by EcoRI, HindIII, SphI, ClaI, Pst and SmaI were compared. The DNAs of 12 different isolates were examined. Clones contained 10-25 kb inserts. When treated with the same restriction endonuclease, the restriction patterns generated from the DNA of each isolate varied. The overall restriction patterns differed from the one obtained from plasmid pYMK3. Many sites which were outside of the transposable element region of plasmid pYMK3 were found to be missing from the clonned DNA. Although plasmid pYMK3 contains a single EcoRI site, several of the clonned inserts contained more than one EcoRI site. These results suggested that while the cloned yeast DNA contained the kanamycin and ura3 genes of the transposable element of plasmid pYMK3, they did not contain other regions of that plasmid. Thus only the transposable element region, but not the plasmid region of plasmid pYMK3 had been cloned.

An isolate of the above described clones, designated as pMKK79-b, was found to contain four EcoRI fragments of 10, 5.0, 4.0, and 0.5 kb in size. The 4.0 kb fragment was identified as the plasmid pBR329. In order to determine whether the three other EcoRI fragments were of yeast chromosomal origin, the plasmid DNA of pMKK79-b was digested with EcoRI and the 10, 5.0 and 0.5 kb fragments were purified, nick-translated with ^{32}p-alpha-dATP and used as probes in a DNA-DNA hybridization. Total DNA of yeast strain YF436 (trp1$^-$ ura3$^-$, cir^o) was blotted to nitrocellulose paper. The DNA fragments from YF436 were then hybridized to the denatured radiolabeled probe. The results of the hybridization are shown in Figure 6. All three fragments were found to be capable of hybridizing two homologous fragments of yeast chromosomal DNA. These results therefore suggest that an association of the transposable element of plasmid pYMK3 with the yeast chromosomal DNA had occurred.



Example 7

Eukaryotic Transposition Using ars-Based Plasmids

The plasmids described previously in the examples as being useful in obtaining transposition in eukaryotes i.e. pYMK2, pYMK3 and their source vector pYCDE2 comprise the 2 yeast origin of replication. Like the 2 -based yeast plasmids, ars-based plasmids are commonly used as cloning vectors in yeast. Both the 2 and ars (autonomously replicating sequence) regions function as replicons in yeast. Both are useful in the present invention but the ars-based plasmids are preferred because they are less stable in the yeast host, a property which can facilitate detection of a transposition event, and because the lower copy number of the ars region in the yeast host decreases the occurrence of homologous recombination.

To confirm the utility of ars-based plasmids in the present invention, plasmid pYMK12 was constructed and tested. To prepare pYMK12, a 1.4 kb EcoRI fragment containing the ars region and coding for trp was extracted from plasmid Yrp7 (see Tschumper and Carbon, 1980, Gene, 10, 157-166) and cloned into the EcoRI site of pBR322. Into the resulting construct was cloned the EcoRISalI fragment from pYMK2 described previously when encodes the ADHI promoter, the nuclear localization sequence of the SV40 T antigen, the Tn5 transposase structural gene and the CYC1 terminator, thereby creating pYMK12 shown in Figure 14. The plasmid pYMK12

represents a useful vector for introducing a transposable element into yeast in a manner substantially as described for pYMK2 but having an ars-replication sequence rather than a 2 replication region.

Transposable element Tn5/ura was introduced within the Ampicillin resistance coding region of pYMK12, providing pYMK12/Tn5/ura. In parallel, a control vector was created from pYMK12 by extracting a SalI/XhoI fragment (shown as "Del/pYMK13 ", in Figure 14) from the 3' end of the transposase gene of pYMK12, thereby providing pYMK13 into which the Tn5/ura region was inserted as described for pYMK12. Thus, pYMK13/Tn5/ura provides a construct which, apart from its inability to encode a functional mRNA transcript of the transposase gene, has substantially the same characteristics as pYMK12/Tn5/ura.

In separate experiments, cells of yeast strains CMY135 were transformed with pYMK12/Tn5/ura and pYMK13/Tn5/ura, respectively, using procedures outlined in Example 4. Colonies surviving on uracil were screened further for the trp phenotype. Of the 3,000 colonies screened, 0.05% of those resulting from pYMK12/Tn5/ura were ura^+ trp^- whereas none of the pYMK13/Tn5/ura (the deletion derivative) transformants exhibited this phenotype; all of the latter were ura^+ trp^- indicating that transposition occurred only in the pYMK12/Tn5/ura transformants.

Subsequent analysis of total DNA of a *ura*$^+$ *trp*$^-$ colony revealed that plasmid pYMK12/Tn5/*ura* was no longer present in the cell as expected given the poor stability of ars-based plasmids in yeast. Further analysis provided a 20 kb EcoRI genomic fragment to which probes for ars, *ura*, neomycin and transposase hybridized but to which probes for trp and pBR322 did not (some very weak homology) indicating that the Tn5/*ura* element had become genomically incorporated substantially independently of plasmid DNA incorporation. The isolated 20 kb genomic fragment appeared to be in a single copy although it is slightly larger than the Tn5/*ura* element *per se*. Mapping of the 20 kb fragment using a variety of restriction enzymes and probes revealed that the non-Tn5/*ura* fragments on the 20 kb fragment were not plasmid-borne suggesting that the Tn5/*ura* element *per se* had transposed from its plasmid vector.

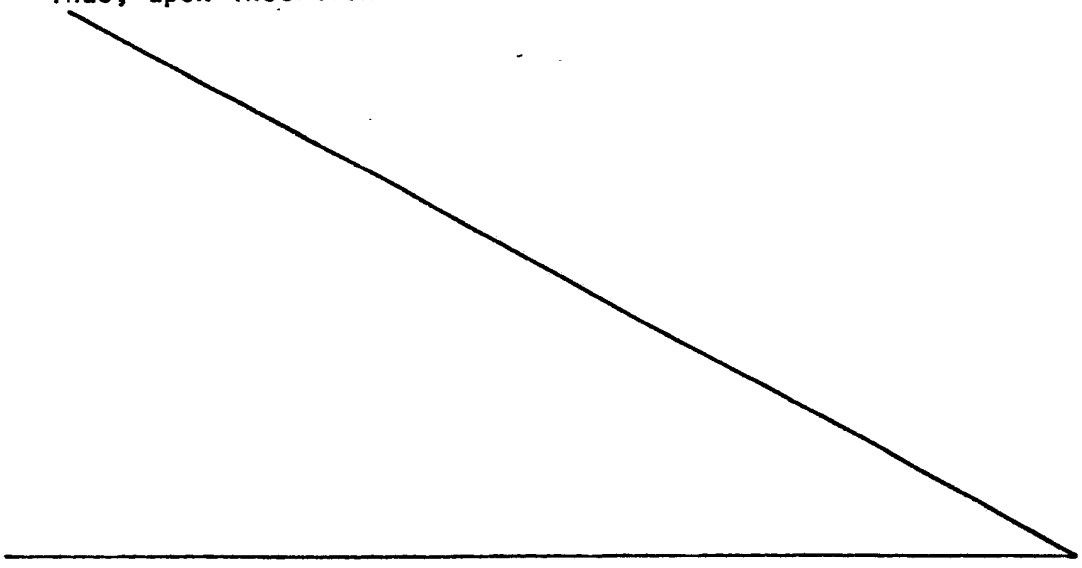
Accordingly, transposition of the yeast genome can be accomplished using vectors which contain a 2μ replication origin or, preferably, an ars sequence in order to introduce the transposable element into the host according to the present invention.

SUBSTITUTE SHEET

Example 8

Construction of a Plasmid Containing a Transposable Element Capable of Controlling Gene Expression

Plasmid pYMK100 was constructed in order to provide a transposable element which, when integrated into chromosomal DNA, would be capable of controlling the expression of those chromosomal genes adjacent to the site of its insertion. The distinguishing features of plasmid pYMK100 are that it contains a transposable element into which a heterologous promoter has been inserted. Importantly, the promoter is capable of directing transcription toward and past one of the ends of the transposable element. Thus, upon insertion



of the transposable element into a chromosomal site, the heterologous promoter will become operably linked to the chromosomal DNA sequence adjacent to the insertion site. The heterologous promoter of plasmid pYMK100 is the gal 10 promoter of Saccharomyces cerevisiae (St. John, P.P., et al., J. Mol. Biol., 152:285-316 (1981); Johnston, M., et al., Mol. Cell. Biol., 4:1440-1448 (1984); Guarente, L., et al., Proc. Natl. Acad. Sci. USA, 79:7410-7414 (1982); Fried, H.M., et al., Mol. Cell. Biol., 5:99-108 (1985)).

Plasmid pYMK100 was constructed as follows. Plasmid YCDE2 was incubated in the presence of SphI and KpnI restriction endonucleases in order to liberate a 6.85 kb SphI-KpnI fragment containing the two micron origin of replication, the plasmid pBR322 origin of replication, the yeast trp1 gene and the ampicillin resistance determinant of plasmid pBR322. The 6.85 kb SphI-KpnI fragment was isolated and purified.

A derivative of plasmid pYMK2 which contained a transposable element was purified and designated pYMK4. This plasmid is essentially identical to that of plasmid pYMK3 except that the insertion site of the transposable element in plasmid pYMK4 is somewhat further from the two micron origin of replication than is the transposable element of plasmid pYMK3. In addition, the orientations of the transposable elements in plasmids pYMK3 and plasmid pYMK4 are inverted with respect to one another. Plasmid pYMK4 was incubated in the presence of SphI restriction endonuclease, and a 4.5 kb SphI-SphI fragment was isolated. This fragment contained the Tn5 transposase gene operably

linked to a nuclear localization signal sequence region which was itself operably linked to the yeast ADHI promoter. In addition, this fragment contained the left hand inverted terminal repeat region of the Tn5 transposable element.

Plasmid pYMK4 was additionally incubated in the presence of SphI and BamHI restriction endonucleases, in order to isolate a 2.2 kb fragment having SphI and BamHI termini. This 2.2 kb fragment contained the yeast ura3 gene.

The gal 10 promoter of Saccharomyces cerevisiae was isolated on an approximately 685 bp fragment having EcoRI and BamHI termini.

In order for the transposable element of plasmid pYMK100 to be capable of undergoing transposition, it was necessary that it have both left and right handed inverted repeated termini regions. An oligonucleotide was synthesized which contained a transposase binding site, and which could function as the right hand inverted repeated region necessary for transposition. The transposase recognition site of the oligonucleotide was bracketed by synthetically derived EcoRI and KpnI restriction endonuclease cleavage sites. The nucelotide sequence of the synthetic oligonucleotide is shown below:

5' AATTCGTGTATAAGAGTCAGGTAC 3'
5' TTAAGCACATATTCTCAGTCCATG 5'

The EcoRI recognition site is located at the left end of the above oligonucleotide; the KpnI restriction endonuclease site is located at the right hand

termini. The transposase recognition site lies between these two restriction endonuclease cleavage sites.

Plasmid pYMK100 was therefore produced by ligating the SphI end of the 6.85 kb SphI-KpnI fragment (from plasmid pYCDE2) to the SphI end nearest to the ADHI promoter region of the 4.5 kb SphI-SphI fragment of plasmid pYMK4. The 2.2 kb SphI-BamHI fragment (of plasmid pYMK4) was ligated to the free SphI end (adjacent to the inverted left terminal repeated region) of the SphI-SphI fragment. The EcoRI-BamHI fragment containing the gal 10 promoter sequence region was ligated to the free BamHI site. Finally, the plasmid molecule was completed through the ligation of the synthetic EcoRI-KpnI transposase recognition sequence fragment to the free KpnI site (of the fragment from plasmid pYCDE2) and to the free EcoRI site of the gal 10-containing fragment. Figure 7 shows a restriction endonuclease cleavage map of plasmid pYMK100.

Thus, plasmid pYMK100 contains a transposase gene capable of expressing a transposase enzyme in a eukaryotic cell. The plasmid also contains a transposable element which is recognized by the transposase enzyme, and which can undergo transposition into DNA present in the nucleous of a eukaryotic cell. When such transposition occurs, the gal 10 promoter will direct the transcription of DNA sequences which (through transposition) are now located adjacent to the synthetic right hand terminal region of the transposable element. Thus, the transcription of such DNA sequences will be dependant upon the induction (by galactose) of the gal 10 promoter. The consequence,

therefore, of transposition next to a particular DNA sequence, is to render the transcription of that DNA sequence under the control of the conditional gal 10 promoter.

Variations to the constructs herein exemplified which improve the frequency of transposition and/or allow for control over transposition are also envisioned by the present invention. Control over transposition can be attained, for example, by linking an inducible heterologous promoter such as the promoters of the melibiase gene, gal 1 gene or gal 10 gene of S. cerevisiae, to the transposase gene of Tn5 or Mu for example. By regulating the presence or concentration of inducer such as galactose in the culture medium, the expression of the transposase gene and, by consequence, the transposition event, can be controlled. Such a modification can be essential in those cells which cannot tolerate constitutive expression of transposase and the resulting high frequency of transposition events.

Eukaryotic cells typically package DNA by shrouding it in protein known as histones. This can be a disadvantage if plasmid vector DNA and the transposable element contained on it become shrouded in histones, particularly when the sites within the transposable element termini to which transposase binds are histone coated and therefore possibly rendered inert to transposase action. While the embodiments exemplified herein clearly establish that transposition can be obtained within eukaryotes notwithstanding the ability of the host to histone-package DNA, it is believed that the frequency of transposition in these hosts may be enhanced by incorporating within the transposable element termini

and adjacent the transposase binding sites therein, a DNA sequence which is relatively inert to histone binding. Such sequences are known in the art as upstream activating sequences (uas) (see Ginger et al., 1985 Cell, 40, 767-774 and references cited therein).

One example of a plasmid in which a transposable element having uas sequence located appropriately therein is pYMK30. The plasmid YMK30 comprises, as a transposable element, truncated (25bp) Tn5 termini each bordering the uas sequence of gal of S. cerevisiae and, between the uas sequences, the ura coding region. This transposable element has been assembled on p329 liberated and then cloned into pYMK12 described previously. In addition to containing uas sequences adjacent to the Tn5 truncated termini, pYMK30 comprises the heterologous gal 10 promoter linked operably to the Tn5 structural transposase gene. Plasmid pYMK30 is shown in Figure 15 and the DNA sequence of one terminus/uas segment is provided in Figure 16. The other terminus is substantially as indicated in Figure 16 but is expressed in the reverse direction. Thus, plasmid YMK30 combines two advantageous modifications; the incorporation of sequences to which histones do not bind, located at sites sufficiently close to transposase binding sites as to interfere with histone binding at those sites; and an inducible heterologous promoter operably linked to the Tn5 transposase structural gene. Transformation of host cells with such vectors is performed substantially as hereinbefore described and transposed hosts similarly isolated.

SUBSTITUTE SHEET

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of invention following, in general, the principles of the invention and including such dispartures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.

International Application No: PCT/GB 87/00598

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 17, line 30 of the description [1]

A. IDENTIFICATION OF DEPOSIT [2]

Further deposits are identified on an additional sheet [3] ☑

Name of depositary institution [4]

AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution (including postal code and country) [4]

12301 PARKLAWN DRIVE, ROCKVILLE, MARYLAND 20852 U.S.A.

Date of deposit [5]	Accession Number [6]
19/6/1986 - 19 JUNE 1986	ATCC 20800

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet ☐

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [8] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the international application when filed (to be checked by the receiving Office)

E Smith

(Authorized Officer) E. SMITH ROOM 5/6 INTERNATIONAL UNIT. EXT. 6906

☐ The date of receipt (from the applicant) by the International Bureau [10] was

(Authorized Officer)

International Application No: PCT/GB 87/00598

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 18, line 2 of the description [1]

A. IDENTIFICATION OF DEPOSIT [3]

Further deposits are identified on an additional sheet ☑ [3]

Name of depositary institution [4]

AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution (including postal code and country) [4]

12301 PARKLAWN DRIVE, ROCKVILLE, MARYLAND 20852 U.S.A

Date of deposit [5]	Accession Number [6]
1/8/86 · 1 AUGUST 1986	ATCC 67173

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet ☐

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [8] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the international application when filed (to be checked by the receiving Office)

E Smith

(Authorized Officer) E. SMITH ROOM 5/6

INTERNATIONAL UNIT. EXT. 6906

☐ The date of receipt (from the applicant) by the International Bureau [10]

was

(Authorized Officer)

International Application No: PCT/GB 87/00598.

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 18, line 2 of the description [1]

A. IDENTIFICATION OF DEPOSIT [2]

Further deposits are identified on an additional sheet ☑ [3]

Name of depositary institution [4]

AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution (including postal code and country) [4]

12301 PARKLAWN DRIVE, ROCKVILLE
MARYLAND 20852, U.S.A.

Date of deposit [5]	Accession Number [6]
11/8/86 11 AUGUST 1986	ATCC 67185

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet ☐

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [9] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the international application when filed (to be checked by the receiving Office)

E Smith.
(Authorized Officer) E. SMITH ROOM 5/6
INTERNATIONAL UNIT. EXT. 6906

☐ The date of receipt (from the applicant) by the International Bureau [10]

wse

(Authorized Officer)

Claims:

1. A recombinant DNA molecule which comprises a transposase gene having a restriction endonuclease recognition site, such that said transposase gene is capable of being operably linked to a DNA sequence selected from the group consisting of: a heterologous promoter region sequence or a nuclear localization signal sequence.

2. A DNA construct which comprises a transposase gene operably linked to a nuclear localization signal sequence.

3. A DNA construct which comprises a transposase gene operably linked to a heterologous promoter region.

4. The DNA construct of claim 3 wherein said promoter region is capable of directing the expression of said transposase gene in a prokaryote.

5. The DNA construct of claim 3 wherein said promoter region is capable of directing the expression of said transposase gene in a eukaryote.

6. The DNA construct of claim 5 wherein said transposase gene is additionally operably linked to a nuclear localization signal sequence.

7. The DNA construct of any claims 3-6 wherein a restriction endonuclease recognition site is present between said transposase gene and said heterologous promoter region.

8. The DNA construct of claim 6 wherein a restriction endoclease recognition site is present between said nuclear localization signal sequence and said transposase gene.

9. A transposable element which contains an exogenous DNA sequence.

10. The transposable element of claim 9 wherein said exogenous DNA sequence comprises a detectable marker gene.

11. The transposable element of claim 10 wherein said detectable marker gene is selected from the group consisting of: genes which confer cellular resistance to antibiotics, the luciferase gene, genes involved in biosynthesis, and genes involved in metabolism.

12. The transposable element of any of claims 9-11 wherein said exogenous DNA sequence comprises a DNA construct which comprises a transposase gene operably linked to a heterologous promoter region.

13. The transposable element of claim 12 wherein said transposase gene is additionally operably linked to a nuclear localization signal sequence.

14. The transposable element of claim 12 wherein said transposase gene is capable of expressing a transposase enzyme, said enzyme being capable of catalyzing the transposition of said transposable element.

15. The transposable element of claim 13 wherein said transposase gene is capable of expressing a transposase enzyme, said enzyme being capable of catalyzing the transposition of said transposable element.

16. The transposable element of claim 9 wherein said exogenous DNA sequence contains a heterologous promoter region sequence.

17. The transposable element of claim 15 wherein said heterologous promoter region sequence is capable of directing the transcription of an additional DNA sequence when said additional DNA sequence is linked to said transposable element.

18. The transposable element of claim 17 wherein said heterologous promoter region sequence is operably linked to a secretory signal sequence.

19. The transposable element of claim 16 wherein said heterologous promoter region is capable of directing gene expression in a prokaryote.

20. The transposable element of claim 16 wherein said heterologous promoter region is capable of directing gene expression in a eukaryote.

21. The transposable element of any of claims 16-20 wherein said heterologous promoter region sequence is a conditional promoter region sequence.

22. The transposable element of claim 21 wherein said transposable element additionally contains and expressible gene whose product is capable of (i) recognizing said conditional promoter region sequence and (ii) controlling the ability of said conditional promoter region sequence to direct gene expression.

23. A DNA molecule which is comprised of the transposable element of claim 21 and an expressible gene whose product is capable of (i) recognizing said conditional promoter region sequence and (ii) controlling the ability of said conditional promoter region to direct gene expression.

24. A DNA molecule which comprises the transposable element of claim 16 and DNA construct, said DNA construct comprising a

transposase gene, said transposase gene being operably linked to a heterologous promoter region sequence.

25. A DNA molecule which comprises a transposable element and a DNA construct, said DNA construct comprising a transposase gene, said transposase gene being operably linked to a heterologous promoter region sequence.

26. The DNA molecule of claim 25 wherein said transposase gene is capable of expressing a transposase enzyme, said enzyme being capable of catalyzing the transposition of said transposable element.

27. A DNA molecule which comprises a transposable element and a DNA construct, said DNA construct comprising a transposase gene, said transposase gene being operably linked to a nuclear localization signal sequence.

28. The DNA molecule of claim 27 wherein said nuclear localization signal sequence is additionally operably linked to a heterologous promoter region sequence.

29. The DNA molecule of claim 29 wherein said exogenous DNA sequence comprises a detectable marker gene.

30. The DNA molecule of claim 29 wherein said exogenous DNA sequence comprises a detectable marker gene.

31. The DNA molecule of claim 30 wherein said detectable marker gene is selected from the group consisting of: genes which confer cellular resistance to antibiotics, the luciferase gene, genes involved in biosynthesis, and genes involved in metabolism.

32. A method for inducing transposition in a prokaryotic cell which comprises:

(a) providing to said prokaryotic cell,

(i) a transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a heterologous promoter region, wherein said transposase gene expresses a transposase enzyme capable of recognizing said transposable element (i) and directing its transposition in said prokaryotic cell, and

(b) permitting said DNA construct (ii) to express said transposase gene and direct the transposition of said transposable element (i).

33. The method of claim 32 wherein said transposable element (i) is incapable of directing its own transposition.

34. The method of claim 32 wherein said heterologous promoter region is derived from said prokaryotic cell

35. The method of claim 33 wherein said heterologous promoter region is derived from said prokaryotic cell.

36. The method of any of claims 32-35 wherein said transposable element contains an exogenous DNA sequence.

37. The method of any of claims 32-35 wherein said transposable element and said DNA construct are present on the same DNA molecule.

38. The method of claim 36 wherein said transposable element and said DNA construct are present on the same DNA molecule.

39. A method for inducing genetic transposable in a eukaryotic cell which comprises:

(a) providing to said eukaryotic cell,

(i) a transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a nuclear localization signal sequence, said nuclear localization signal sequence being operably linked to a promoter region; said transposase gene being heterologous to said promoter

region or to said nuclear localization signal sequence, wherein said promoter region directs the synthesis of a transposase enzyme, said enzyme being linked to the amino acid sequence encoded by said nuclear localization signal sequence, and capable of entering the nucleus of said eukaryotic cell; said enzyme being capable of recognizing said transposable element (i) and directing its transposition in said eukaryotic cell, and

(b) permitting said DNA construct (ii) to express said transposase enzyme and direct the transposition of said transposase element (i).

40. The method of claim 39 wherein said transposable element (i) is incapable of directing its own transposition.

41. The method of claim 39 wherein said promoter region is derived from said eukaryotic cell.

42. The method of claim 40 wherein said promoter region is derived from said eukaryotic cell.

43. The method of any of claims 39-41 wherein said transposable element contains an exogenous DNA sequence.

44. The method of any of claims 39-42 wherein said transposable element and said DNA construct are present on the same DNA molecule.

45. The method of claim 43 wherein said transposable element and said DNA construct are present on the same DNA molecule.

46. A method for controlling the expression of a target gene in a prokaryotic cell which comprises:

(a) providing to said prokaryotic cell,

(i) a transposable element, said transposable element containing an exogeneous DNA sequence which comprises a conditional and heterologous promoter region sequence capable of directing the transcription in said prokaryotic cell of an additional DNA sequence when said additional DNA sequence is linked to said transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a heterologous promoter region, wherein said transposase gene expresses a transposase enzyme capable of recognizing said transposable element (i) and directing its transposition in said prokaryotic cell,

(b) permitting said DNA construct (ii) to express said transposase gene and direct the transposition of said transposable element (i) and

(c) examining for said prokaryotic cell in which said transposition has occurred and in which the expression of said target gene is regulated by said conditional and heterologous promoter.

47. The method of claim 46 wherein said prokaryotic cell is additionally provided with an expressible gene whose product is capable of (i) recognizing said conditional and heterologous promoter region sequence and (ii) controlling the ability of said conditional promoter region sequence to direct gene expression.

48. A method for controlling the expression of a target gene in a eukaryotic cell which comprises:

(a) providing to said eukaryotic cell,

(i) a transposable element, said transposable element containing an exogenous DNA sequence which comprises a conditional and heterologous promoter region sequence capable of directing the transcription in said eukaryotic cell of an additional DNA sequence when said additional DNA sequence is linked to said transposable element, and

(ii) a DNA construct which comprises a transposase gene operably linked to a nuclear localization signal sequence, said nuclear localization signal sequence being operably linked to a promoter region; said

transposase gene being heterologous to said promoter region or to said nuclear localization signal sequence, wherein said promoter region directs the synthesis of a transposase enzyme, said enzyme being linked to the amino acid sequence encoded by said nuclear localization signal sequence, and capable of entering the nucleus of said eukaryotic cell; said enzyme being capable of recognizing said transposable element (Pi) and directing its transposition in said eukaryotic cell, and

(b) permitting said DNA construct (ii) to express said transposase enzyme and direct the transposition of a said transposable element (i) and

(c) examining for said eukaryotic cell in which said transposition has occurred and in which the expression of said target gene is regulated by said conditional and heterologous promoter.

49. The method of claim 48 wherein said eukaryotic cell is additionally provided with an expressible gene whose product is capable of (i) recognizing said conditional and heterologous promoter region sequence (ii) controlling the ability of said conditional promoter region sequence to direct gene expression.

50. The plasmid pMKK58, and its functional derivatives.

51. The plasmid pYMK3, and its functional derivatives.

52. The plasmid pYMK100, and its functional derivatives.

53. A DNA construct which comprises a transposase gene operably linked to a heterologous promoter wherein said promoter functions in a Streptomyces host.

54. The DNA construct of claim 53 which comprises a transposable element.

55. A Streptomyces host having a transposable element within genomic DNA thereof.

56. A DNA construct comprising a transposable element having upstream activating sequences incorporated adjacent to the termini of said transposable element.

57. A DNA molecule comprising the construct of claim 56 and a second DNA construct comprising a transposase gene operably linked to a heterologous promoter and a nuclear localization signal sequence.

FIGURE 1

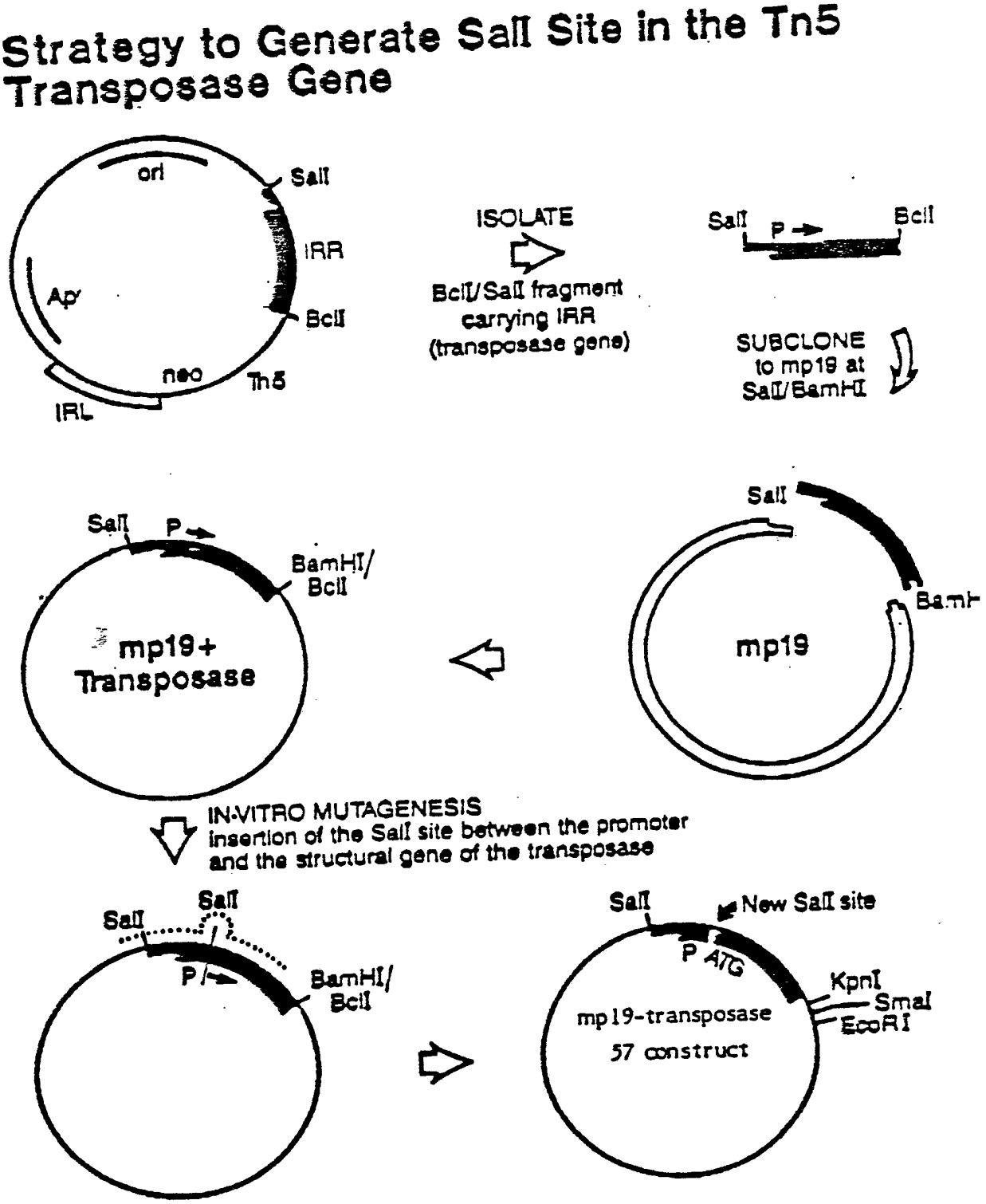

Strategy to Generate SalI Site in the Tn5 Transposase Gene
ori
SalI
IRR
Ap
BclI
neo
Tn5
IRL
ISOLATE
BclI/SalI fragment carrying IRR (transposase gene)
SalI
P
BclI
SUBCLONE to mp19 at SalI/BamHI
SalI
P
BamHI/ BclI
mp19+ Transposase
SalI
BamH
mp19
IN-VITRO MUTAGENESIS
insertion of the SalI site between the promoter and the structural gene of the transposase
SalI
SalI
P
BamHI/ BclI
SalI
New SalI site
P ATG
KpnI
SmaI
EcoRI
mp19-transposase
57 construct

FIGURE 2

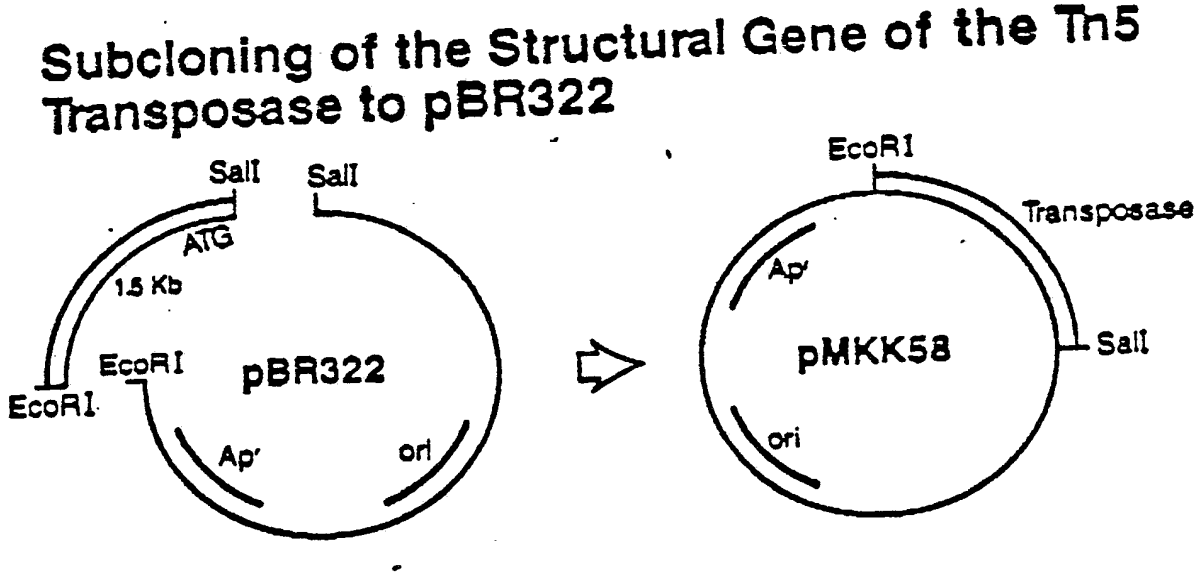

Subcloning of the Structural Gene of the Tn5 Transposase to pBR322
SalI
SalI
ATG
1.5 Kb
EcoRI
EcoRI
pBR322
Ap'
ori
EcoRI
Transposase
Ap'
pMKK58
SalI
ori

FIGURE 3

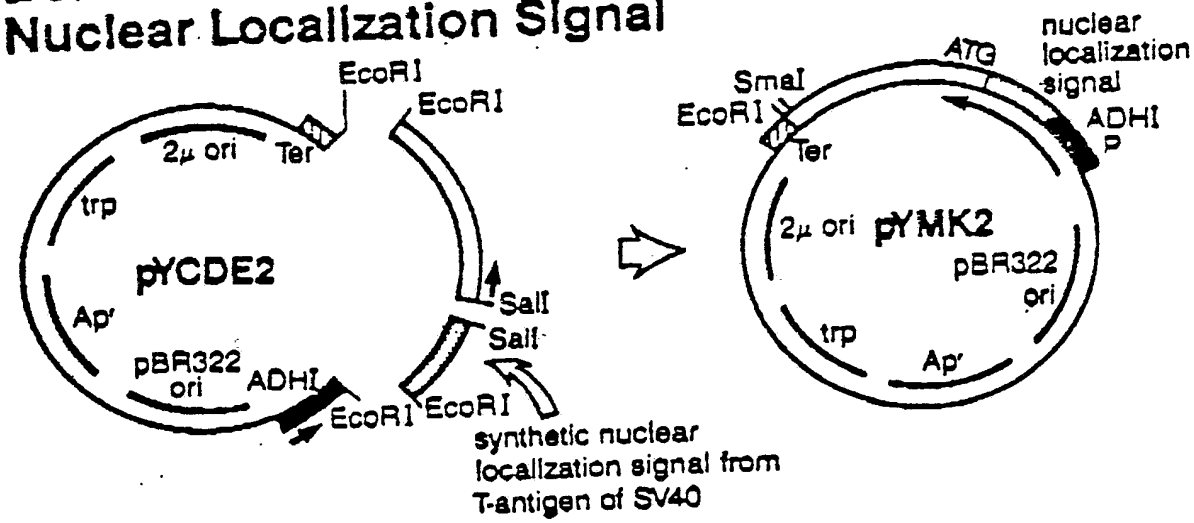
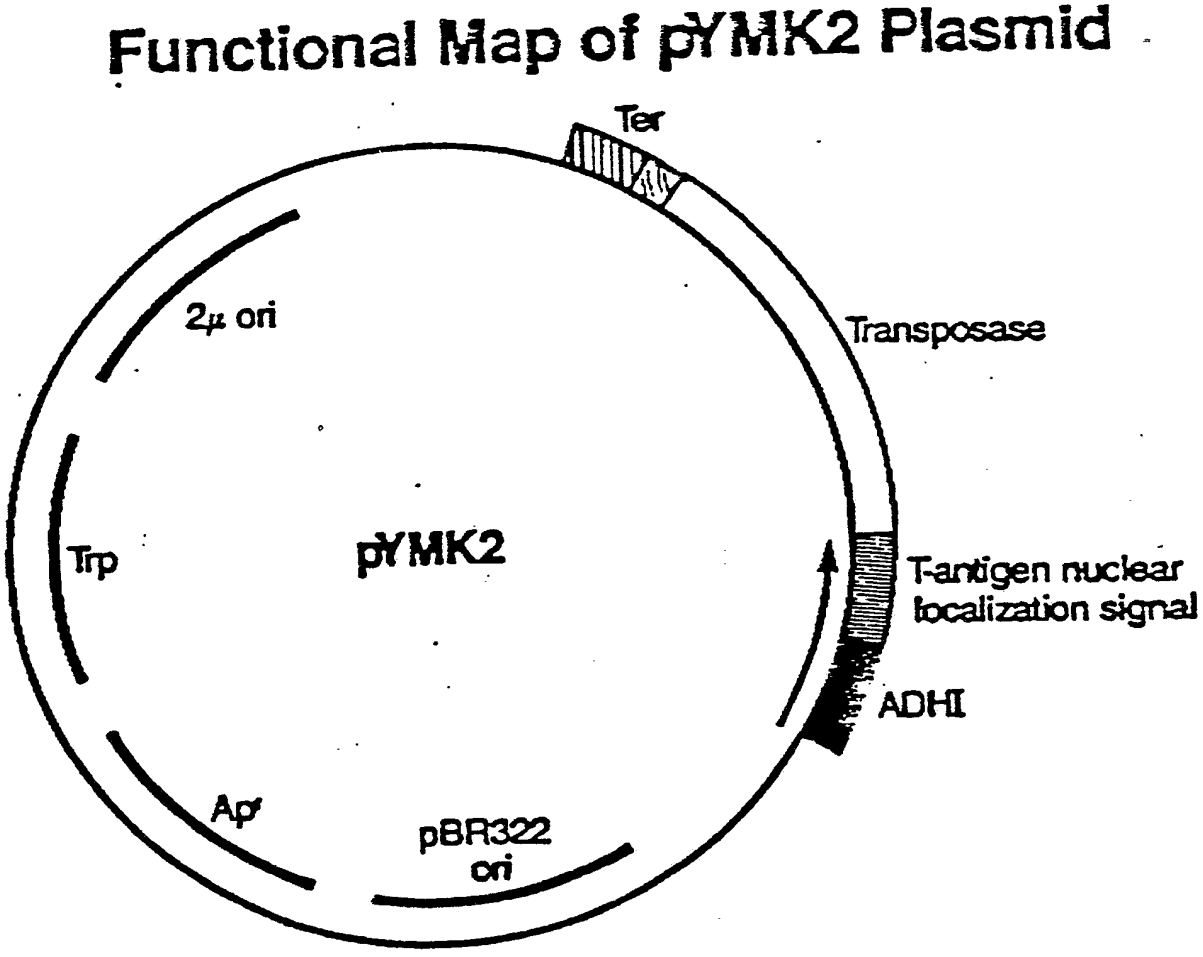

Cloning of the Transposase Structural Gene Behind the ADHI Promoter and the T-antigen Nuclear Localization Signal
EcoRI
EcoRI
2μ ori
Ter
trp
pYCDE2
Ap'
pBR322 ori
ADHI
EcoRI
EcoRI
SalI
SalI
synthetic nuclear localization signal from T-antigen of SV40
SmaI
EcoRI
ATG
nuclear localization signal
ADHI P
Ter
2μ ori
pYMK2
pBR322 ori
trp
Ap'

FIGURE 4


Functional Map of pYMK2 Plasmid
Ter
2μ ori
Transposase
Trp
pYMK2
T-antigen nuclear localization signal
ADHI
Ap
pBR322 ori

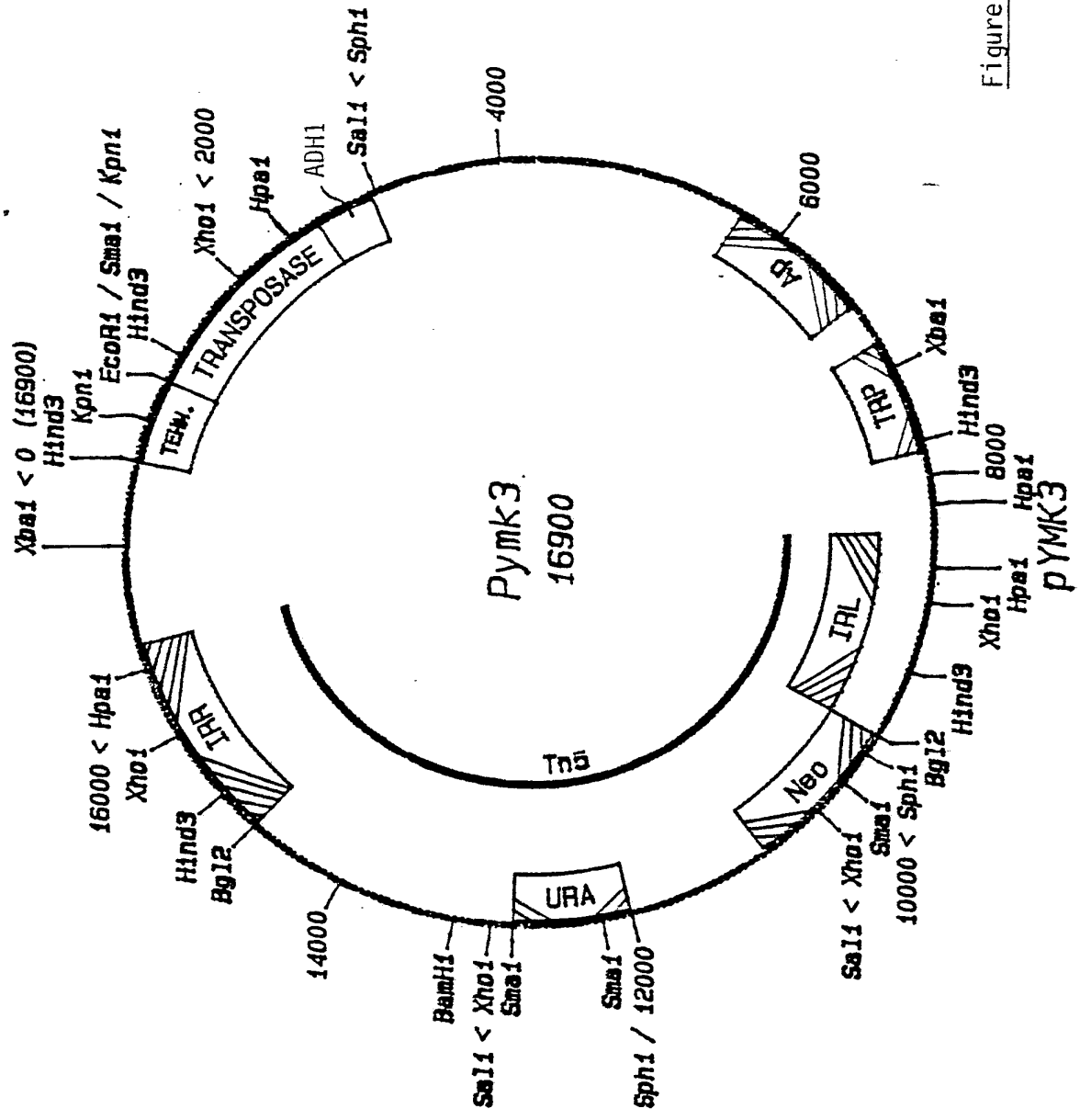
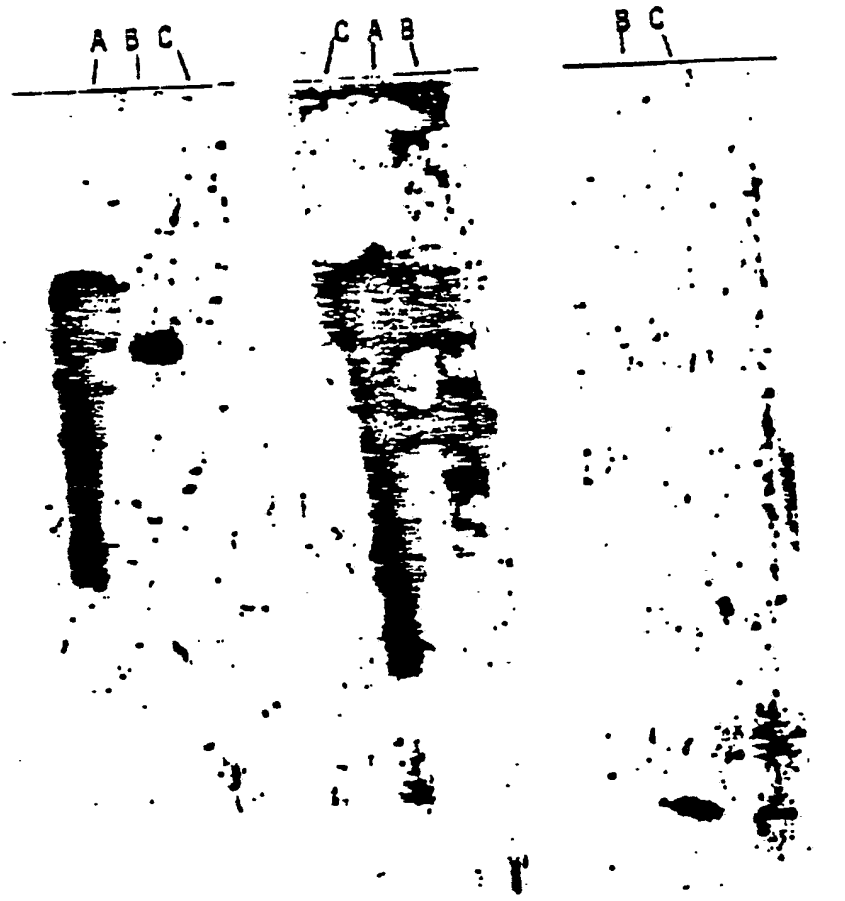

Pymk3
16900
Tn5
4000
6000
8000
10000 < Sph1
Sal1 < Xho1
Sma1
Bgl2
Hind3
Xho1
Hpa1
Hpa1
Xba1
Hind3
AR
TRP
IRL
Neo
URA
Sma1
Sph1 / 12000
BamH1
Sal1 < Xho1
Sma1
14000
Hind3
Bgl2
16000 < Hpa1
Xho1
IRR
Xba1 < 0 (16900)
Hind3
Kpn1
EcoR1 / Sma1 / Kpn1
Hind3
TRANSPOSASE
TEMP.
Xho1 < 2000
Hpa1
ADH1
Sal1 < Sph1
pYMK3

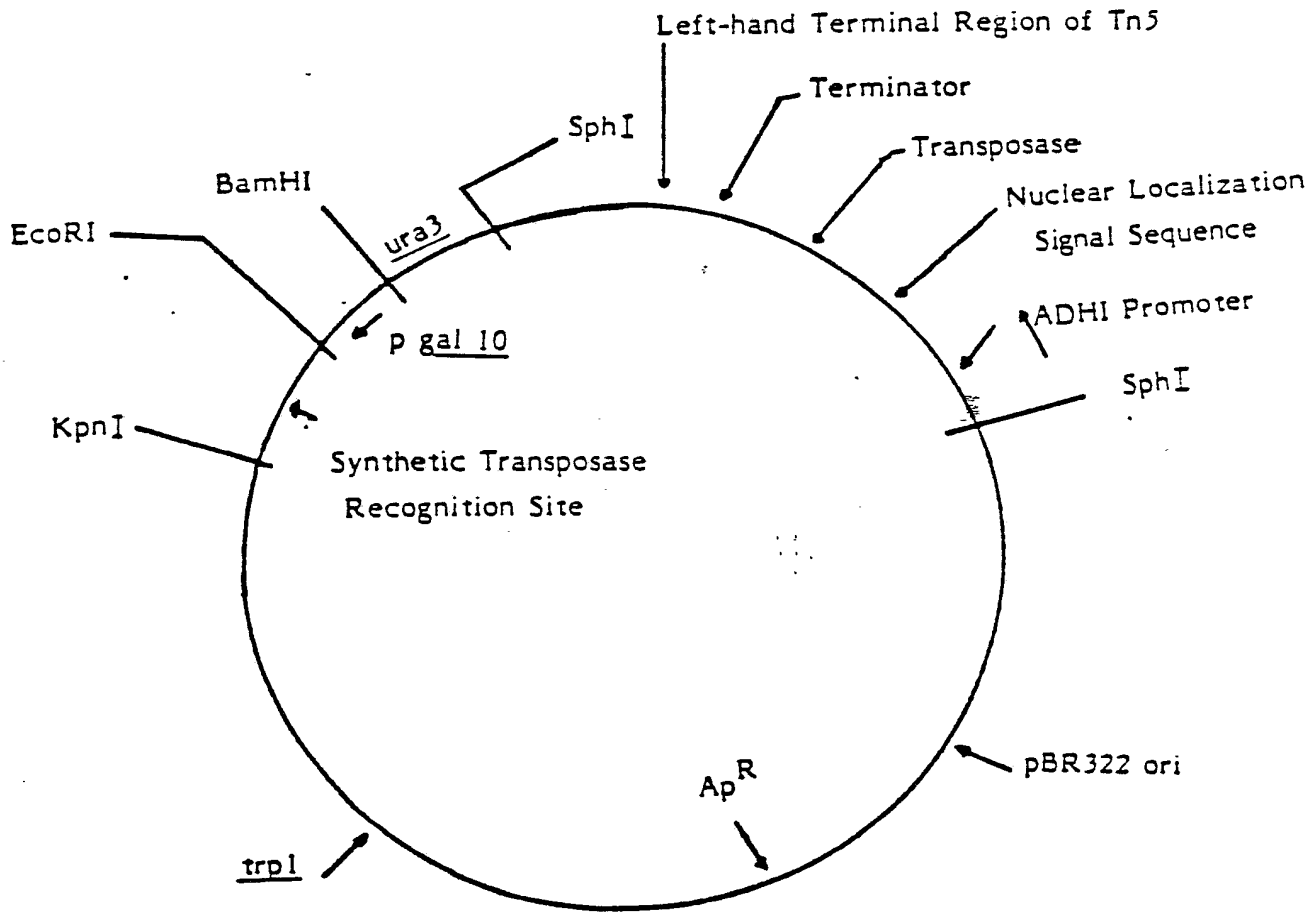
Figure 5

FIGURE 6



A - λ HIND III
B - Total DNA from Y*436/pYMK3 cut with EcoRI
C - Total DNA from Y*436 cut with EcoRI

FIGURE 7



pYMK 100

The Use of an Engineered Transposable Element to Yield Conditional Mutants

Legend

A-B genes

M marker gene

P natural promoter

Tn element end (transposase binding site)

P controllable promoter/ direction of transcription

T terminator

ORF open reading frame - 3 versions provide 3 different reading frames

stop stop codon (translational stop)

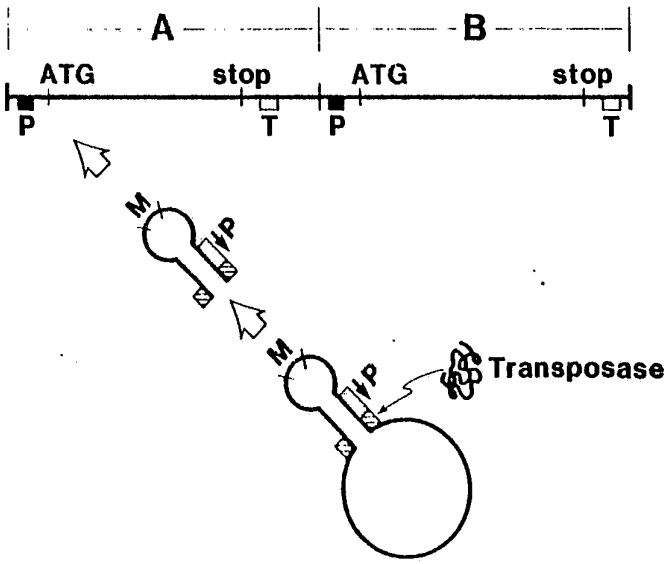


Figure 8

Mode I Tn carried controllable promoter substitutes for the natural promoter: coding region A is unchanged.

Mode II Tn carried controllable promoter substitutes for the natural promoter **And** Tn carries open reading frame (dependant on controllable promoter for expression) is fused to a gene A reading frame in frame.

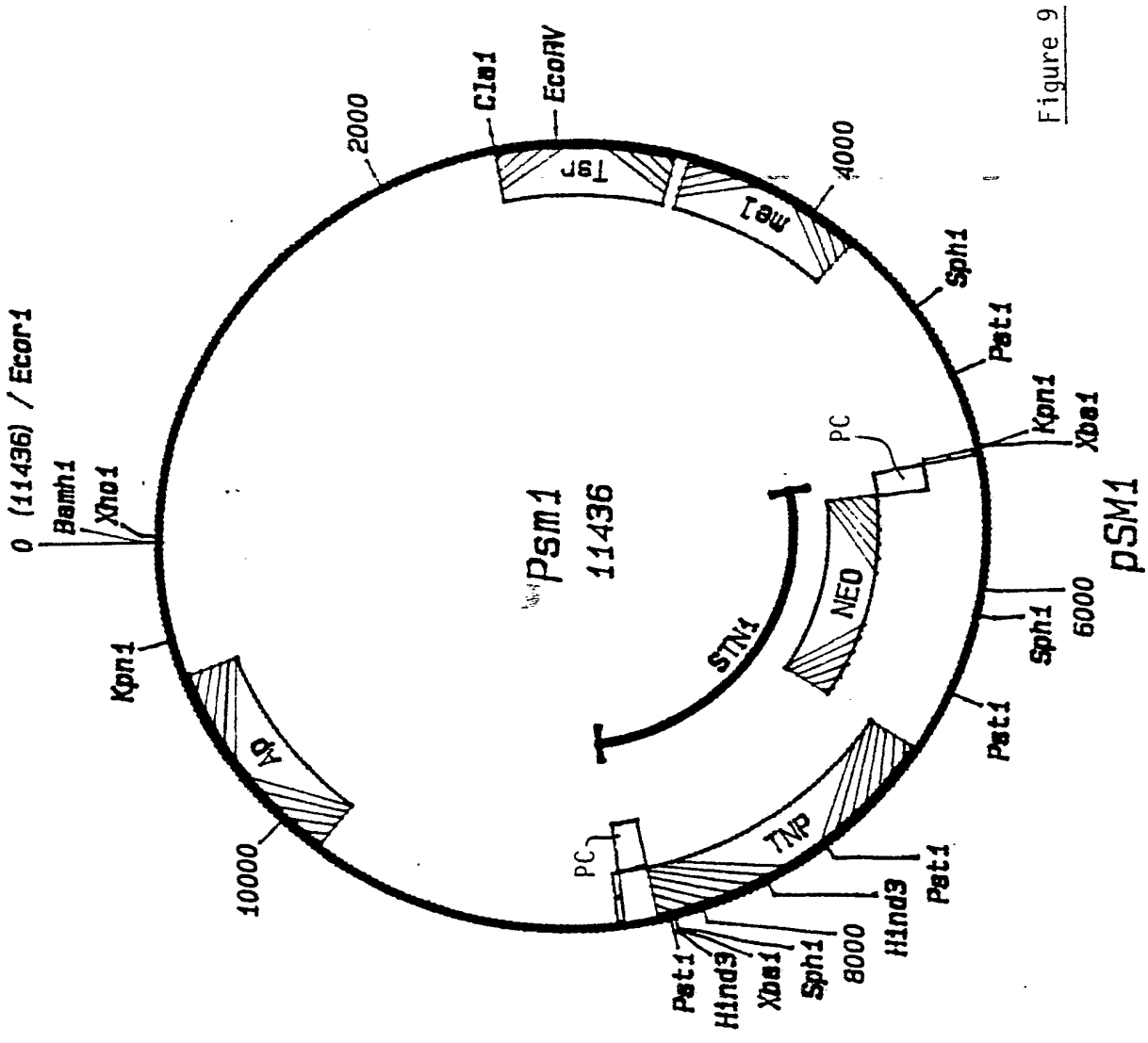

0 (11436) / Ecor1
Bamh1
Xho1
Kpn1
Ap
10000
2000
Cla1
EcoRV
Tar
Tet
4000
Sph1
Pst1
Kpn1
Xba1
PC
NEO
STN1
Psm1
11436
pSM1
6000
Sph1
Pst1
TNP
Hind3
Pst1
8000
Sph1
Xba1
Hind3
Pst1
PC

Figure 9

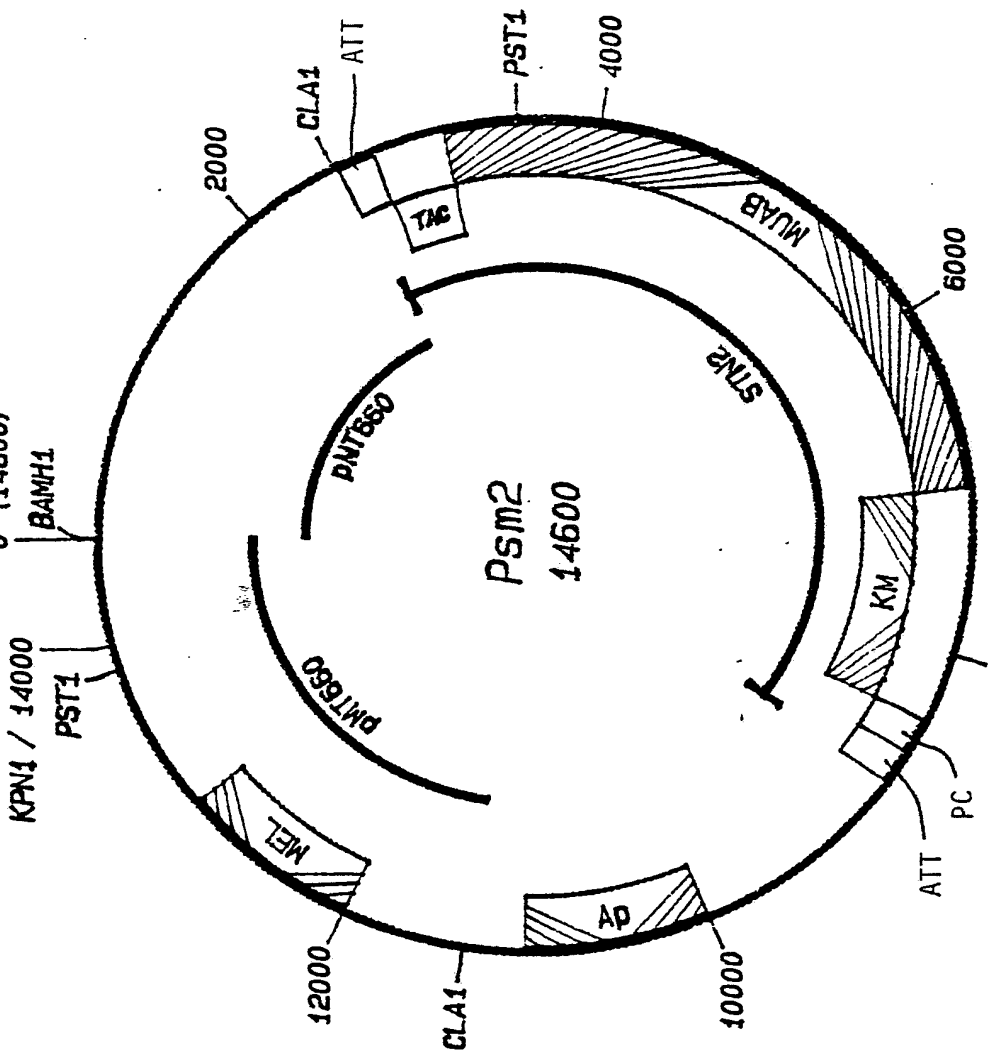


Figure 10

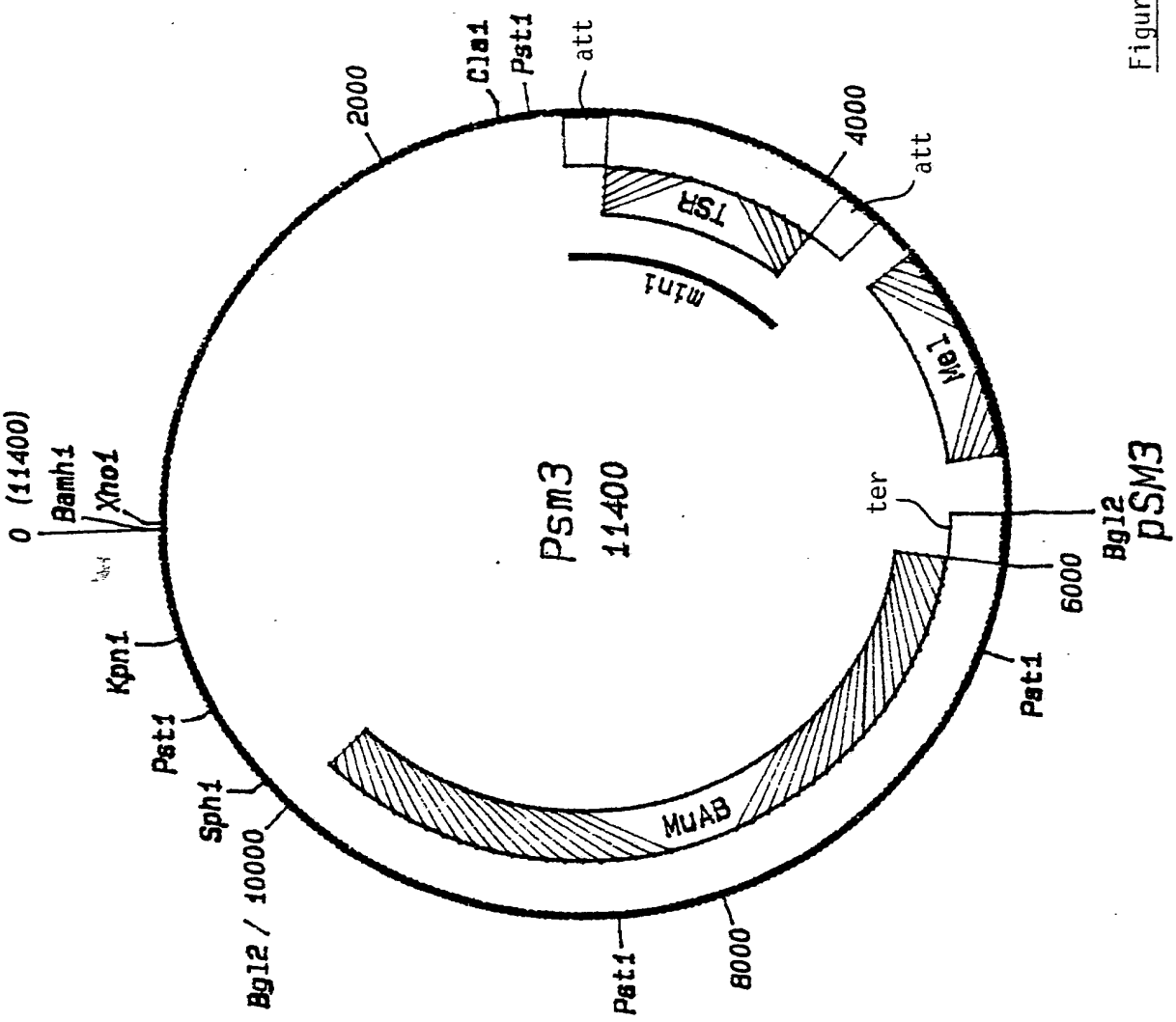

0 (11400)
Bamh1
Xho1
Kpn1
Pst1
Sph1
Bgl2 / 10000
Pst1
8000
Pst1
6000
Bgl2
pSM3
ter
Mel
att
4000
TSR
mini
att
Cla1
Pst1
2000
Psm3
11400
MuAB

Figure 11

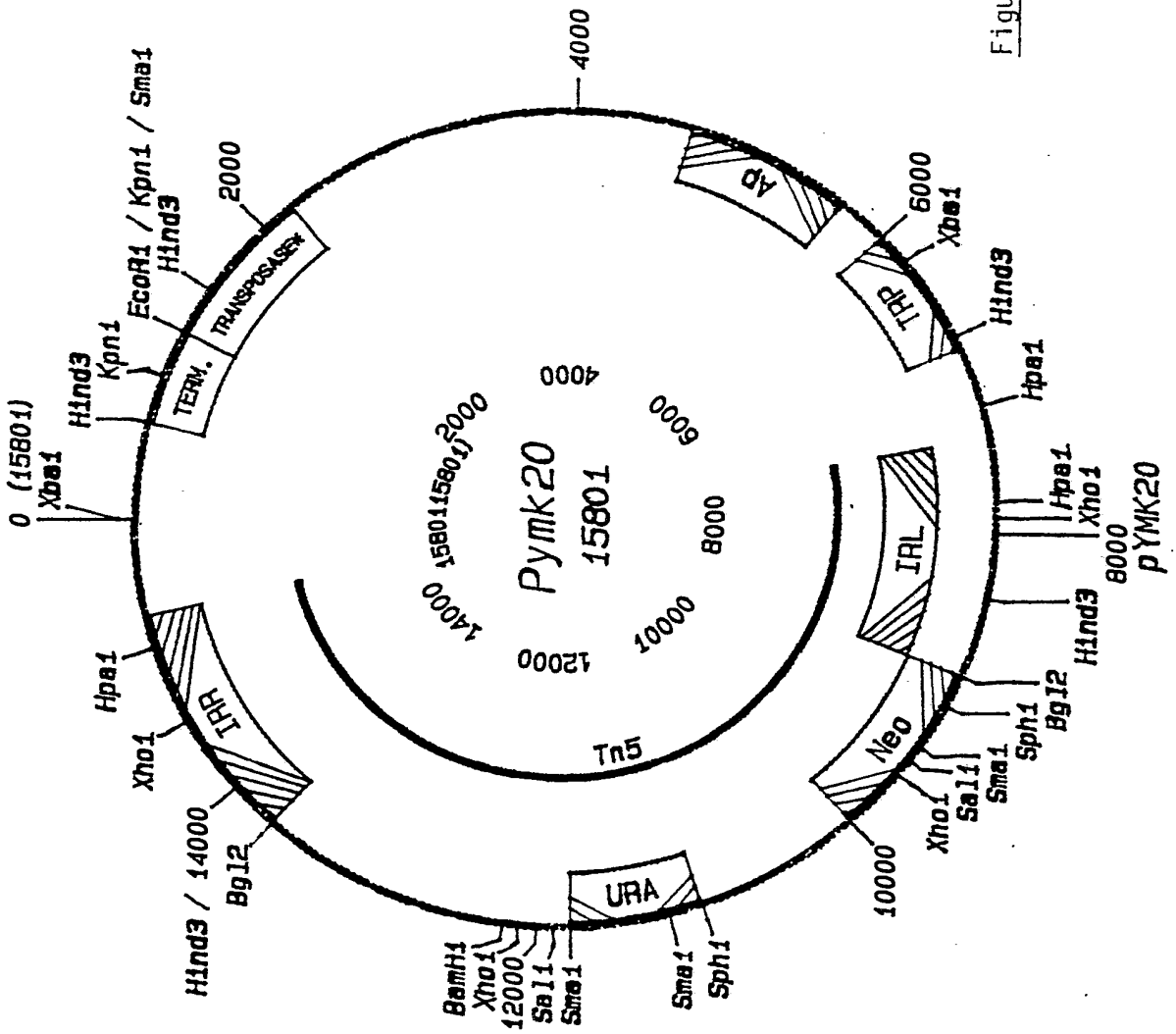

pYMK20
Pymk20
15801
0 (15801)
2000
4000
6000
8000
10000
12000
14000
15801
Xba1
Hind3
Kpn1
EcoR1 / Kpn1 / Sma1
Hind3
TERM.
TRANSPOSASE
Hpa1
Xho1
Bg12
Hind3 / 14000
TRA
TRA
Hpa1
Xba1
Hind3
Hpa1
Xho1
8000
Hind3
Bg12
Sph1
Sma1
Sal1
Xho1
10000
Neo
IRL
Tn5
12000
BamH1
Xho1
Sal1
Sma1
URA
Sma1
Sph1
Ap
TRP
Xba1
6000

Figure 12

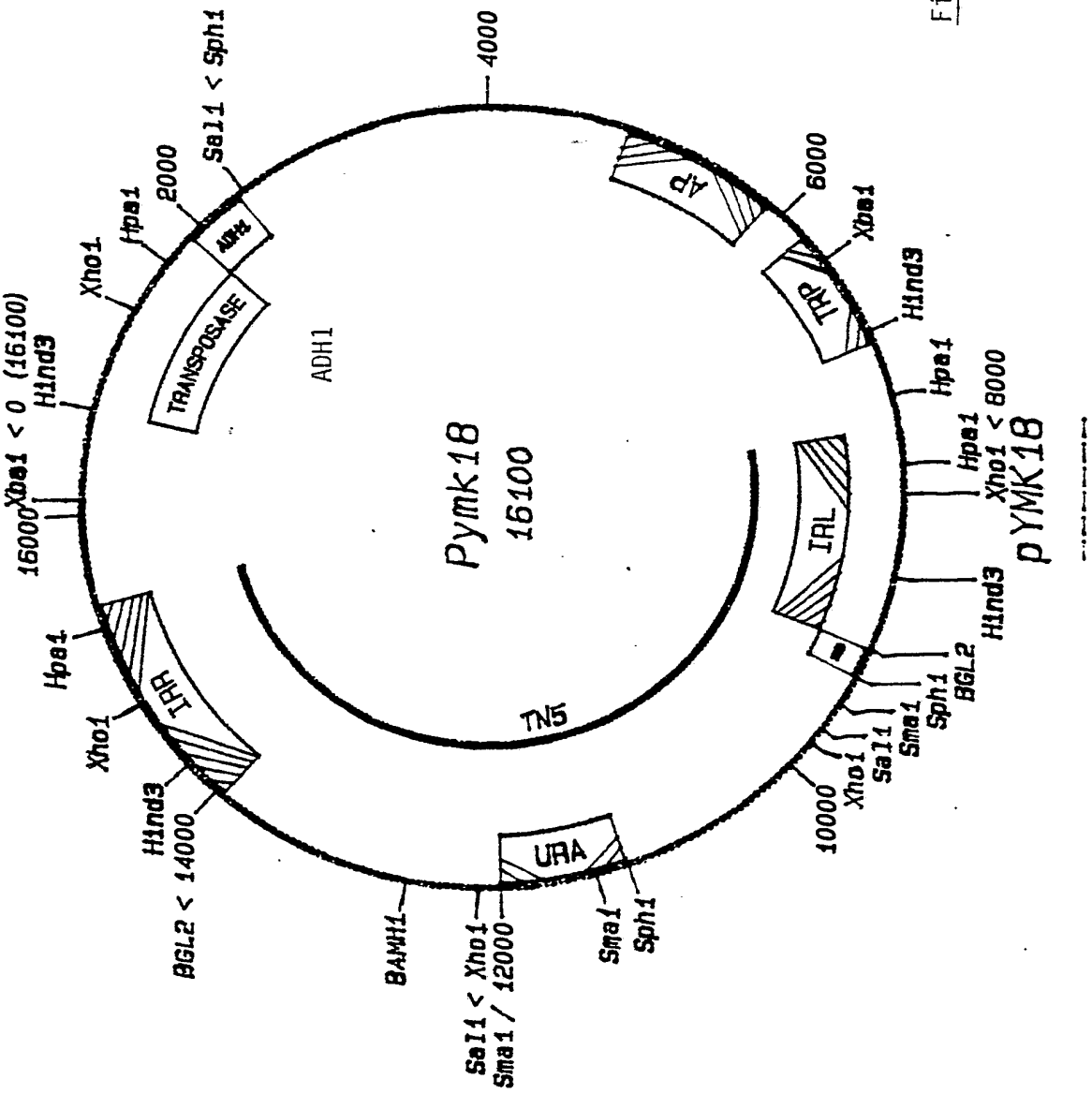

Pymk18
16100
pYMK18
ADH1
TRANSPOSASE
TRP
IRL
TN5
URA
TRP
AMP
4000
6000
8000
10000
12000
14000
16000
2000
Sal1 < Sph1
Hpa1
Xho1
Hind3
Xba1 < 0 (16100)
Xba1
Hpa1 < 8000
Hpa1
Xho1
Hind3
BGL2
Sph1
Sma1
Sal1
Xho1
Hind3
Hpa1
Xho1
Hind3
BGL2 < 14000
BAMH1
Sal1 < Xho1
Sma1 / 12000
Sma1
Sph1

Figure 13

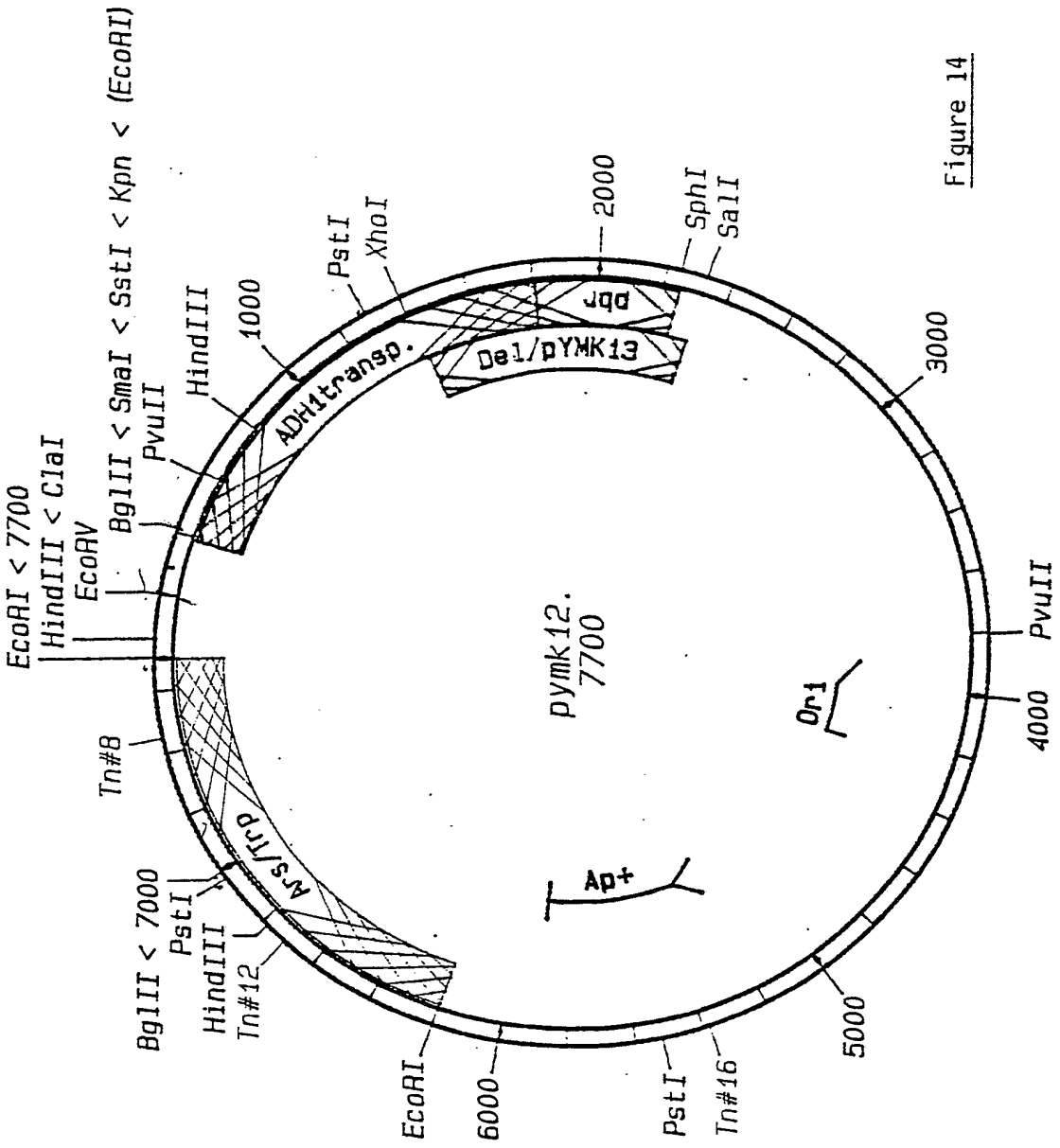

EcoRI < 7700
HindIII < ClaI
EcoAV
BglII < SmaI < SstI < KpnI < (EcoRI)
PvuII
HindIII
1000
PstI
XhoI
2000
SphI
SalI
3000
PvuII
4000
5000
PstI
Tn#16
6000
EcoRI
BglII < 7000
PstI
HindIII
Tn#12
Ars/Trp
Tn#8
pymk12.
7700
Ori
Ap+
ADH1transp.
Del/pYMK13
pbr

Figure 14

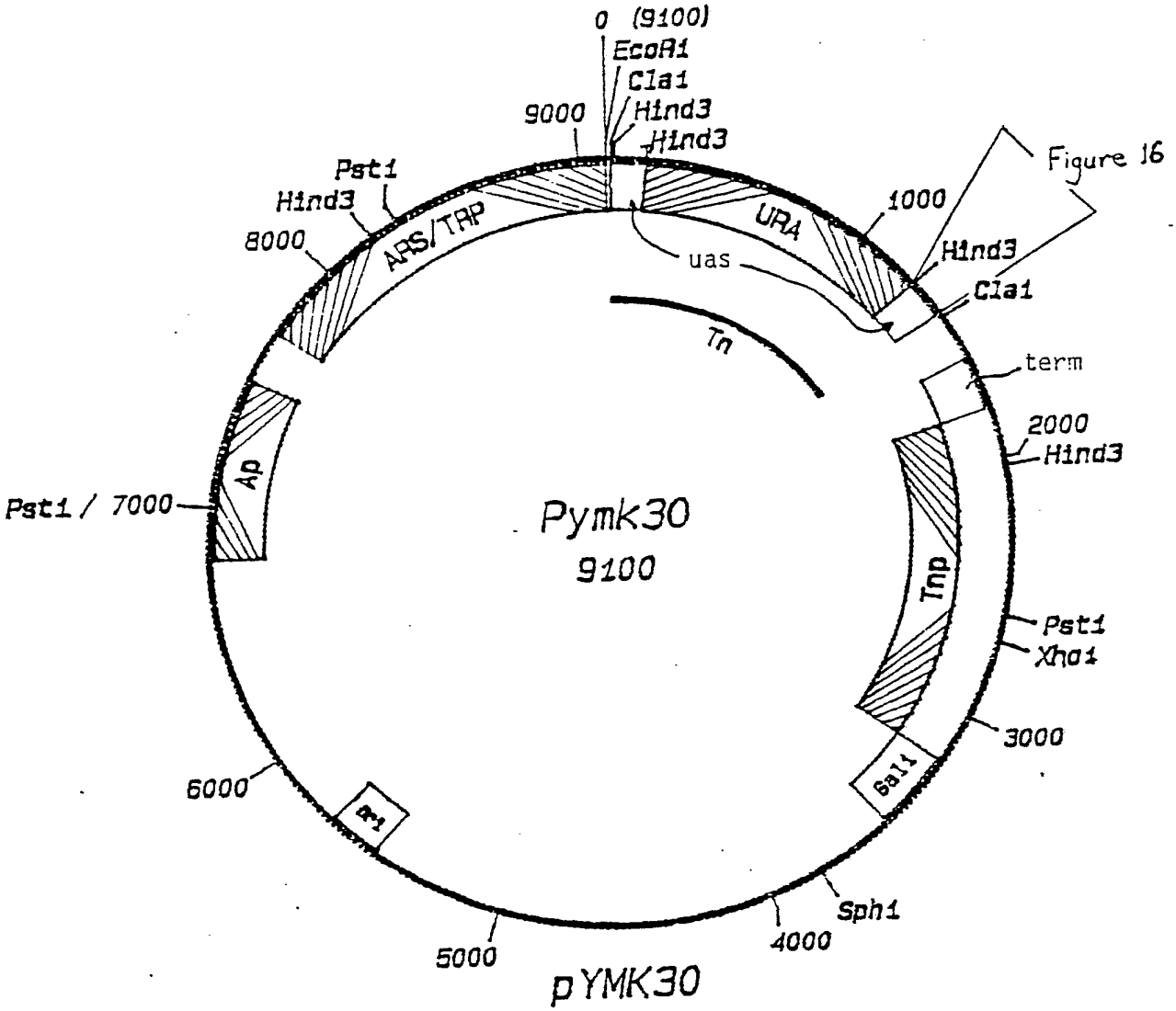


Figure 15

Figure 16

```
Sal1 Cla1  Sca1                          Tn5 inverted repeat
TCGACATCGATAGTACTGCTGACTCTTATACACAAGTAGCGTCCTGAACGGAACCTTT-
   GTAGCTATCATGACGACTGAGAATATGTGTTCATCGCAGGACTTGCCTTGGAAA-

CCCGTTTTAGATCTCGGAAGACTCTCCTCCGAAGCTTAGATCTG
GGGCAAAATCTAGAGCCTTCTGAGAGGAGGCTTCGAATCTAGACAGCT
      BglII      -uas-          HindIII     SalI
                                    BglII
```

Figure 16

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 87/00598

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00; C 12 N 1/20; (C 12 N 1/20, C 12 R 1:465)

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X	US, A, 4670388 (G.M. RUBIN et al.) 2 June 1987 see the whole document	1-49
A	EP, A, 0130047 (AGRIGENETICS RESEARCH ASSOCIATES LTD.) 2 January 1985 see pages 20-22 and 30-33	1-49
A	EP, A, 0091723 (IMPERIAL CHEMICAL INDUSTRIES PLC) 19 October 1983 see the whole document	9-49
X	Chemical Abstracts, vol. 99, 1983 (Columbus, Ohio, US) G. Sermonti et al.: "Properties of transposon SCTn1 of Streptomyces coelicolor A3(2)", see page 171, abstract no. 116940e & MGG, Mol. Gen. Genet. 1983, 191(1), 158-61	55

---

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
29th November 1987	20 JAN 1988
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	P.C.G. VAN DER PUTTEN

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 8700598

SA 18452

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 22/12/87
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A- 4670388	02-06-87	None	
EP-A- 0130047	02-01-85	None	
EP-A- 0091723	19-10-83	JP-A- 58152899	10-09-83
		US-A- 4590162	20-05-86